



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 105983**

**TO: Susan Ungar**  
**Location: CM1/8E12/8B05**  
**Art Unit: 1642**  
**Wednesday, October 15, 2003**

**Case Serial Number: 09/854124**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**CM1-6B02**  
**Phone: 305-9203**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Ungar,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**THIS PAGE BLANK (USPTO)**

Hart, Edward

105983

**From:** Ungar, Susan  
**Sent:** Monday, October 13, 2003 2:23 PM  
**To:** Chan, Christina  
**Cc:** Hart, Edward  
**Subject:** Rush sequence search for 09/854,124

Hi

I need a rush sequence search for 09/854,124, SEQ ID NOS 5, 6, 7 - these are polynucleotide fragments.

Ed Hart has agreed to submit this search for me, please send authorization directly to him.

Thanks  
Susan Ungar  
1642  
703-305-2181  
CM1-8B05

10/14/03

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

ABSSD4  
3-NA

10/15/03

**THIS PAGE BLANK (USPTO)**



Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .rnpn and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn

**Because they contain data that is confidential, the results of Pending database searches should not be left in the case .**



# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art found, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art not found:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ Des



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:00:56 ; Search time 1477 Seconds  
(without alignments)  
8309.335 Million cell updates/sec

Title: US-09-854-124-6  
Perfect score: 300  
Sequence: 1 aattccgtgtgtgtgtgtga.....cgaaaggagaacattttac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rtd.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	284.8	94.9	858	6	BD156103	BD156103 Primer fo
2	284.8	94.9	858	9	AK001275	AK001275 Homo sapi
3	284.8	94.9	1174	9	AF151880	AF151880 Homo sapi
4	284.8	94.9	1180	9	BC010164	BC010164 Homo sapi
5	284.8	94.9	1188	9	AF201939	AF201939 Homo sapi
6	284.8	94.9	1203	9	BC003186	BC003186 Homo sapi
7	284.8	94.9	3097	9	AK091519	AK091519 Homo sapi
8	283	94.3	667	6	BD147616	BD147616 Primer fo
9	227.6	75.9	116792	9	HS1100H13	AL035419 Human DNA
10	223.4	74.5	454	6	AX336730	AX336730 Sequence
11	167.4	55.8	933	5	AB097169	AB097169 xenopus l
12	125.2	41.7	160356	9	AC009068	AC009068 Homo sapi
13	125.2	41.7	185321	9	AC123908	AC123908 Homo sapi
14	120	40.0	245210	2	AC137771	AC137771 Homo sapi
15	97	32.3	207307	2	AC117085	AC117085 rattus no
16	97	32.3	228373	2	AC118833	AC118833 Rattus no
17	93.8	31.3	188292	2	AC103360	AC103360 Mus muscu
18	85.2	28.4	825	3	AY071234	AY071234 Drosophill
19	85.2	28.4	72722	3	AC004373	AC004373 Drosophill
20	85.2	28.4	132666	3	AC005270	AC005270 Drosophill
21	85.2	28.4	165158	2	AC141687	AC141687 Apis mell
22	85.2	28.4	182897	3	AC092232	AC092232 Drosophill
23	85.2	28.4	218794	2	AC018084	AC018084 Drosophill
24	85.2	28.4	292393	3	AE003576	AE003576 Drosophill
25	59.2	19.7	33103	3	CEF31C3	Z92784 Caenorhabdi
26	55.4	18.5	4593	3	LEINACETYL	M96635 Leishmania
27	55.4	18.5	110000	2	LMFLCH36_16	Continuation (17 o
28	51	17.0	245210	2	AC137771	AC137771 Homo sapi
29	47.6	15.9	139399	8	AP002865	AP002865 Oryza sat
30	47.6	15.9	182756	8	AC007789	AC007789 Oryza sat
31	47.2	15.7	37949	8	SPBC725	AL034352 S.pombe c
32	40.6	13.5	190459	2	AP005419	AP005419 Oryza sat
33	39.4	13.1	21024	8	AP002055	AP002055 Arabidops
34	39.4	13.1	105768	8	AC069474	AC069474 Arabidops
35	38.6	12.9	82804	3	BMMIF1	AL606837 Brugia ma
36	38.4	12.8	310550	1	SC0939113	AL939113 Streptomy
37	38	12.7	10596	1	AE004746	AE004746 Pseudomon
38	38	12.7	88223	2	AC084755	AC084755 Homo sapi
39	38	12.7	241891	2	AC132675	AC132675 Rattus no
40	38	12.7	259329	2	AC109570	AC109570 Rattus no
41	37.2	12.4	298450	1	SC0939107	AL939107 Streptomy
42	37	12.3	228940	2	AC126890	AC126890 Rattus no
43	36.8	12.3	7954	1	TAQ19223	Y19223 Thermus aqu
44	36.8	12.3	10869	1	AE005069	AE005069 Halobacte
45	36.4	12.1	3008	6	AX683109	AX683109 Sequence

ALIGNMENTS

RESULT 1  
BD156103  
LOCUS BD156103 858 bp DNA linear PAT 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD156103  
VERSION BD156103.1 GI:27861861  
KEYWORDS JP 2002191363-A/10946.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 858)  
AUTHORS Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,  
Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.  
TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 10946 09-JUL-2002;  
HELIK RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
PN JP 2002191363-A/10946  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OYA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU  
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
PI KEIICHI NAGAI,TETSUJI OTSUKI  
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
10, C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS (37)..(591).  
Location/Qualifiers  
source 1..858  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 227 a 214 c 222 g 195 t  
ORIGIN  
Query Match 94.9%; Score 284.8; DB 6; Length 858;  
Best Local Similarity 99.3%; Pred. No. 2e-64;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 GTCGGTGAGGCTCTGGCTGCAGCTCGCGCGCCATGAGCTCCGAGGTCGAATTCCT 72  
DB 3 GACCGTGAGGCTCTGGCTGCAGCTCGCGCGCCATGAGCTCCGAGGTCGAATTCCT 62  
QY 73 CGCGGAGAAGGAGCTGTTACCAATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT 132  
DB 63 CGCGGAGAAGGAGCTGTTACCAATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT 122  
QY 133 CGGGGGGAGCTGGGGCTTTTACCTCGTGTTCACCGTGGAAAGTCCCTGTGGCTGGC 192  
DB 123 CGGGGGGAGCTGGGGCTTTTACCTCGTGTTCACCGTGGAAAGTCCCTGTGGCTGGC 182  
QY 193 GATTAACCTGAAACAAACAGACAGAATGTGCGCTCTCCCTCCAGAGTGGATGTAGTA 252  
DB 183 GATTAACCTGAAACAAACAGACAGAATGTGCGCTCTCCCTCCAGAGTGGATGTAGTA 242  
QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGAAAGAACTTTTAC 300  
DB 243 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGAAAGAACTTTTAC 290  
RESULT 2  
AK001275 858 bp mRNA linear PRI 01-AUG-2002  
LOCUS Homo sapiens cDNA FLJ10413 fis, clone NT2RP1000063.  
DEFINITION AK001275  
ACCESSION AK001275.1 GI:7022428  
VERSION oligo capping; fis (full insert sequence).  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoliri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 858)  
AUTHORS Isogai,T. and Otsuki,T.

Direct Submission  
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of NEDO International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
FEATURES Location/Qualifiers  
source 1..858  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP1000063"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RP1"  
/note="cloning vector: pUC19FL3-mRNA from NT2 neuronal precursor cells after 48-hours retinoic acid (RA) induction."  
BASE COUNT 227 a 214 c 222 g 195 t  
ORIGIN  
Query Match 94.9%; Score 284.8; DB 9; Length 858;  
Best Local Similarity 99.3%; Pred. No. 2e-64;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 GTCGGTGAGGCTCTGGCTGCAGCTCGCGCGCCATGAGCTCCGAGGTCGAATTCCT 72  
DB 3 GACCGTGAGGCTCTGGCTGCAGCTCGCGCGCCATGAGCTCCGAGGTCGAATTCCT 62  
QY 73 CGCGGAGAAGGAGCTGTTACCAATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT 132  
DB 63 CGCGGAGAAGGAGCTGTTACCAATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT 122  
QY 133 CGGGGGGAGCTGGGGCTTTTACCTCGTGTTCACCGTGGAAAGTCCCTGTGGCTGGC 192  
DB 123 CGGGGGGAGCTGGGGCTTTTACCTCGTGTTCACCGTGGAAAGTCCCTGTGGCTGGC 182  
QY 193 GATTAACCTGAAACAAACAGACAGAATGTGCGCTCTCCCTCCAGAGTGGATGTAGTA 252  
DB 183 GATTAACCTGAAACAAACAGACAGAATGTGCGCTCTCCCTCCAGAGTGGATGTAGTA 242  
QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGAAAGAACTTTTAC 300  
DB 243 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGAAAGAACTTTTAC 290  
RESULT 3  
AF151880 1174 bp mRNA linear PRI 18-MAY-2000  
LOCUS Homo sapiens CGI-122 protein mRNA, complete cds.  
DEFINITION AF151880  
ACCESSION AF151880  
VERSION AF151880.1 GI:4929712  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1174)  
AUTHORS Lal,C.H., Chou,C.Y., Ch'ang,L.Y., Liu,C.S. and Lin,W.

CONTACT: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@hri.nih.gov">nisc_mgc@hri.nih.gov</a> Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tiongson, E.E., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
CLONE DISTRIBUTION: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 27 Row: 1 Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.
FEATURES Location/Qualifiers 1. .1180 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="LocusID:51659" /db_xref="taxon:9606" /clone="MGC:19836 IMAGE:4098007" /tissue_type="Brain, neuroblastoma" /clone_lib="NIH_MGC_19" /lab_host="DH10B-R" /note="Vector: pOTB7" 68. .625 /codon_start=1 /product="HSPC037 protein" /protein_id="AAH10164.1" /db_xref="GI:14603432" /translation="MDAAEVEFLAEKELVTIIPNFDKIIYLLGGDLGPNGLPVEV PLWALNLKQKCRLLPPEWMDVEKLRMDHREKEETFTPMPSYIMELTKLLNH ASDNIPKADEIRTLVKDMDTRIAKLRVSADSFVROQEAHAKLDNLTLMEINTSGTFL TQALNHMYKLRTNLOPLESTQSODF"
BASE COUNT 316 a 292 c 307 g 265 t
ORIGIN Query Match 94.9%; Score 284.8; DB 9; Length 1180; Best Local Similarity 99.3%; Pred. No. 2.1e-64; Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0; QY 13 GTCGGTGAGGCTCTGGCGTGCAGCTCGCGCGCCATGGAGCTGCGAGTGCGAATTCCT 72 Db 34 GACCGTGAGGCTCTGGCGTGCAGCTCGCGCGCCATGGAGCTGCGAGTGCGAATTCCT 93 QY 73 CGCCGAGAGGAGCTGTTACCATTTACCCACTTCAGTCTGCAGAGATCCTCAT 132 Db 94 CGCCGAGAGGAGCTGTTACCATTTACCCACTTCAGTCTGCAGAGATCCTCAT 153 QY 133 CGGGGGGAGCTCGGGGCTTTTAAACCTGTTTACCCGTGGAAGTGCCTGTGGGTGGC 192 Db 154 CGGGGGGAGCTCGGGGCTTTTAAACCTGTTTACCCGTGGAAGTGCCTGTGGGTGGC 213 QY 193 GATTACCTGAAACAAGACAGAAATGTGCCCTGCTCCCTCCAGATGGATGTAGTAGA 252 Db 214 GATTACCTGAAACAAGACAGAAATGTGCCCTGCTCCCTCCAGATGGATGTAGTAGA 273 QY 253 AAAGTTGGAGAGATGAGGATCATGAACGAAAGAGAACTTTTAC 300 Db 274 AAAGTTGGAGAGATGAGGATCATGAACGAAAGAGAACTTTTAC 321
RESULT 5

## RESULT 5

AF201939	1188 bp	mRNA	linear	PRI 20-JUL-2000
LOCUS				
DEFINITION	Homo sapiens DC5 (DC5) mRNA, complete cds.			
ACCESSION	AF201939			
VERSION	AF201939.1	GI:9295181		
KEYWORDS	FLI_CDNA.			
SOURCE	Homo sapiens (human)			
ORGANISM				
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.			
TITLE	Novel genes expressed in human dentritic cell			
JOURNAL	Unpublished			
REFERENCE	2. (bases 1 to 1188)			
AUTHORS	Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-NOV-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China			
FEATURES	Location/Qualifiers			
source	1..1188			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
gene	/cell_type="dentritic cell"			
	1..1188			
	/gene="DC5"			
CDS	77..634			
	/gene="DC5"			
	/codon_start=1			
	/product="DC5"			
	/protein_id="AAF96875.1"			
	/db_xref="GI:9295182"			
	/translation="MDAAVEFLAEKELVTIIPNFSLDKIYLIIGDLGPNPGLPVEV PLWAINLKOROKRLPPEDMDVEKLEMRDHERKEETETPMPSPYIMELTKLLLNH ASDNIPKADRIETLVKMDWTRIAKLRSVDSFVRQAEAAKLNDLIMEINTSGTFL TQALNHMYKLRNLPLESTQSQDF"			
BASE COUNT	314 a	299 c	307 g	268 t
ORIGIN				
Query Match	94.98;	Score 284.8;	DB 9;	Length 1188;
Best Local Similarity	99.38;	Pred. No. 2.1e-64;		
Matches	286;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
QY	13	GTCCGTGAGGCTCTGGCGCTCGAGCTCGCGCGCATGGAGCTGCCGAGGTGCGAATTCCT	72	
Db	43	GACCGTAGGCTCTGGCGCTCGAGCTCGCGCGCATGGAGCTGCCGAGGTGCGAATTCCT	102	
QY	73	CGCGGAGAAGGAGCTGGTTACCATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT	132	
Db	103	CGCGGAGAAGGAGCTGGTTACCATATCCCAACTTCAGTCTGGACAAAGATCTACCTCAT	162	
QY	133	CGGGGGGAGCTGGGGCTTTTAAACCTGGTTTACCCTGGTGAAGTGCCTCTGGCTGGC	192	
Db	163	CGGGGGGAGCTGGGGCTTTTAAACCTGGTTTACCCTGGTGAAGTGCCTCTGGCTGGC	222	
QY	193	GATTAACCTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGATGATGAGA	252	
Db	223	GATTAACCTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGATGATGAGA	282	
QY	253	AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC	300	
Db	283	AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC	330	
RESULT 6				
BC003186	1203 bp	mRNA	linear	PRI 12-JUL-2001
LOCUS				
DEFINITION	Homo sapiens, HSPC037 protein, clone MGC:673 IMAGE:3503268, mRNA, complete cds.			
ACCESSION	BC003186			
VERSION	BC003186.1	GI:13112024		

KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 1203)
TITLE	Strausberg,R.
JOURNAL	Direct Submission
REMARK	Submitted (13-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
FEATURES	Location/Qualifiers
source	1..1203
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/db_xref="GI:13112025"
	/clones="MGC:673 IMAGE:3503268"
	/tissue_type="Brain, neuroblastoma"
	/clone_lib="NIH_MGC_19"
	/lab_host="DH10B-R"
	/note="Vector: pOTB7"
	85..642
	/codon_start=1
	/product="HSPC037 protein"
	/protein_id="AAH03186.1"
	/db_xref="GI:13112025"
	/translation="MDAAVEFLAEKELVTIIPNFSLDKIYLIIGDLGPNPGLPVEV PLWAINLKOROKRLPPEDMDVEKLEMRDHERKEETETPMPSPYIMELTKLLLNH ASDNIPKADRIETLVKMDWTRIAKLRSVDSFVRQAEAAKLNDLIMEINTSGTFL TQALNHMYKLRNLPLESTQSQDF"
BASE COUNT	326 a 300 c 309 g 268 t
ORIGIN	
Query Match	94.98;
Best Local Similarity	99.38;
Matches	286;
	Conservative 0;
	Mismatches 2;
	Indels 0; Gaps 0;
QY	13
Db	51
QY	73
Db	111
QY	133
Db	171

QY 193 GATTAACTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGATGGATGTAGA 252  
 |||||  
 Db 231 GATTAACTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGATGGATGTAGA 290  
 |||||  
 QY 253 AAAGTTGGAGAGATGAGGATCATGAACGAAGAAGAAACTTTTAC 300  
 |||||  
 Db 291 AAAGTTGGAGAGATGAGGATCATGAACGAAGAAGAAACTTTTAC 338  
 |||||

RESULT 7  
 AK091519  
 LOCUS 3097 bp mRNA linear PRI 15-JUL-2002  
 DEFINITION Homo sapiens CDNA FLJ34200 fls, clone FCBF3019663.  
 AK091519  
 VERSION AK091519.1 GI:21749909  
 KEYWORDS oligo capping; fls (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1  
 Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,  
 Ono,Y., Hotuta,T., Hirao,K., Murakawa,K., Takiguchi,S.,  
 Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,  
 Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,  
 Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,  
 Kawai-Hio,Y., Sekine,K., Nishikawa,T., Kimura,K., Matsuo,K.,  
 Nakamura,Y., Sakino,K., Kikuchi,H., Kanda,K., Wagatsuma,M.,  
 Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,  
 Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.  
 NEDO human cDNA sequencing project

Unpublished

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS 2 (bases 1 to 3097)  
 Isogai,T. and Yamamoto,J.

TITLE  
 JOURNAL Direct Submission  
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: HRI and  
 RAB; annotation: HRI and RAB.

FEATURES  
 source  
 1. 3097  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="FCBF3019663"  
 /tissue\_type="brain"  
 /clone\_lib="FCBF3"  
 /dev\_stage="fetal"  
 /note="cloning vector: pME18SFL3"

BASE COUNT  
 ORIGIN 726 a 703 c 786 g 882 t

Query Match 94.9%; Score 284.8; DB 9; Length 3097;  
 Best Local Similarity 99.3%; Pred. No. 2.2e-64;  
 Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGTAGGCTCTGGCTGAGCTGCGCGCCGATGACGCTGCCGAGGTGGAATTCCT 72  
 |||||

Db 1 GACCGTAGAGCTCTGGCTGAGCTGCGCGCCGATGACGCTGCCGAGGTGGAATTCCT 60  
 |||||

QY 73 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTACGTCTGGACAAGATCTACCTCAT 132  
 |||||

Db 61 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTACGTCTGGACAAGATCTACCTCAT 120  
 |||||

QY 133 CGGGGGGACCTGGGGCCCTTTTAACTGCTTTACCCCTGTTACCCCTGGAAGTGGCCCTGTGGCTGGC 192  
 |||||

Db 121 CGGGGGGACCTGGGGCCCTTTTAACTGCTTTACCCCTGGAAGTGGCCCTGTGGCTGGC 180  
 |||||  
 QY 193 GATTAACTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGATGGATGTAGA 252  
 |||||  
 Db 181 GATTAACTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGATGGATGTAGA 240  
 |||||  
 QY 253 AAAGTTGGAGAGATGAGGATCATGAACGAAGAAGAAACTTTTAC 300  
 |||||  
 Db 241 AAAGTTGGAGAGATGAGGATCATGAACGAAGAAGAAACTTTTAC 288  
 |||||

RESULT 8  
 BD147616  
 LOCUS 667 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 AK091519  
 VERSION BD147616  
 KEYWORDS BD147616.1 GI:27853374  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 667)  
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 Primer for synthesizing full-length cDNA and use thereof  
 Patent: JP 2002191363-A 2459 09-JUL-2002;  
 JOURNAL HELIX RESEARCH INSTITUTE

COMMENT  
 OS Homo sapiens (human)  
 PN JP 2002191363-A/2459  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 200280990  
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU  
 PI SAITO,  
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
 PI KEIICHI NAGAI,TETSUJI OTSUKI

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
 10,  
 PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key

FT source  
 1. 667  
 /organism="Homo sapiens"  
 /location/Qualifiers

FT  
 1. 667  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT  
 ORIGIN 171 a 176 c 177 g 137 t 6 others

Query Match 94.3%; Score 283; DB 6; Length 667;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-64;  
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TGAGGCTCTGGCTGAGCTGCGCGCCGATGACGCTGCCGAGGTGGAATTCCTCGCGG 77  
 |||||

Db 7 TGAGGCTCTGGCTGAGCTGCGCGCCGATGACGCTGCCGAGGTGGAATTCCTCGCGG 66  
 |||||

QY 78 AGAGGAGCTGTTTACCATTATCCCAACTTTCAGTCTGGACAAGATCTACCTCATCGGGG 137  
 |||||

Db 67 AGAGGAGCTGTTTACCATTATCCCAACTTTCAGTCTGGACAAGATCTACCTCATCGGGG 126  
 |||||

QY 138 GGGACCTGGGCTTTTAACTGCTTTACCCCTGGAAGTGGCCCTGTGGCGATTA 197  
 |||||

Db 127 GGGACCTGGGCTTTTAACTGCTTTACCCCTGGAAGTGGCCCTGTGGCGATTA 186  
 |||||

QY 198 ACCTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGGATGTAGAAAAGT 257  
 |||||

Db 187 ACCTGAAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGGATGTAGAAAAGT 246  
 |||||

QY 258 TGGAGAAGATGAGGGATCATGACGAAGAAAGAAACTTTTAC 300  
 |||||||  
 Db 247 TGGAGAAGATGAGGGATCATGACGAAGAAAGAAACTTTTAC 289

## RESULT 9

## HS1100H13/c

## LOCUS

## DEFINITION

HS1100H13 116792 bp DNA linear PRI 18-MAR-2001  
 Human DNA sequence from clone RP5-1100H13 on chromosome 20q11.2  
 Contains the 3' end of gene KIAA1219, a putative novel gene, a DC5  
 pseudogene, the gene for a putative RhoGAP domain containing  
 protein, ESTs, STSs, GSSs and two CpG islands, complete sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Submitted (12-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonequery@sanger.ac.uk  
 On Dec 5, 2000 this sequence version replaced gi:10198625.

During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
 on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 20, constructed by the Sanger Centre Chromosome 20  
 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. RP5-1100H13 is from  
 the library RPCI-5 constructed by the group of Pieter de Jong. For  
 further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCVPAC2

This sequence is the entire insert of clone RP5-1100H13 The true  
 left end of clone RP1-12201 is at 47827 in this sequence. The true  
 right end of clone RP5-927M24 is at 16404 in this sequence.

## FEATURES

## source

1. .116792  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="R2PD:RPCI-5" /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="q11.2"  
 /clone="RP5-1100H13"  
 /clone\_lib="RPCI-5"  
 1331. .1465  
 /note="MER58C repeat: matches 68. .89 of consensus"  
 1587. .13349  
 /gene="dj1100H13.1"  
 join(<1587. .1723,4379. .4487,5260. .5338,8641. .8789,  
 9265. .13349)  
 /gene="dj1100H13.1"  
 /product="dj1100H13.1.1 (KIAA1219, isoform 2)"

## repeat\_region

## gene

## mRNA

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

GH09358 and C. elegans D2085.5) isoform 1)"  
 /note="match: cDNAs: Em:AB033045 Em:AL13298  
 match: ESTs: Em:AI190579 Em:AA256927 Em:N36561 Em:AA874351  
 Em:AI281190 Em:AA180094 Em:AA669533 Em:AA511971  
 Em:AI195307 Em:AA215757 Em:AA156777 Em:AA131113  
 Em:AA18776 Em:AI357456 Em:AI379408 Em:N76170 Em:W52362  
 Em:AA024901 Em:AI334294 Em:AA205264 Em:AA215603  
 Em:AA18686 Em:AA024902 Em:AI379162 Em:AA424825  
 Em:AI202970 Em:AI248059 Em:AA648752 Em:AI359832  
 Em:AI366165 Em:AI224862 Em:AI359759 Em:AI92050 Em:R76319  
 Em:AI050034 Em:AI423927 Em:AA024896 Em:R8702 Em:AI066457  
 Em:R72124 Em:AI361138 Em:R81299 Em:AA282907 Em:LA4403  
 Em:AI500502 Em:Z98499 Em:R83793 Em:AA412128 Em:AA179405  
 Em:AA459722 Em:AA173335 Em:AA283091 Em:H26506 Em:H02652  
 Em:AA258043 Em:N64761 Em:AA255436 Em:AI271673 Em:AI017681  
 Em:AA495863 Em:AI221344 Em:R65916 Em:W07865 Em:AA383163  
 Em:H26520 Em:AA872554 Em:AA343687 Em:AA889796 Em:N68842  
 Em:AA580081 Em:AA156671 Em:AA552080 Em:AA024895"  
 /evidence="not\_experimental"  
 join(<1587. .1723,4379. .4487,5260. .5338,8641. .8789,  
 9265. .9458)  
 /gene="dj1100H13.1"  
 /note="Continues in Em:AL049868 as dj927M24.2  
 match: proteins: Tr:Q9Y3G9 Tr:Q9W2X2 Tr:Q9U4G6 Tr:Q18431"  
 /codon\_start=3  
 /evidence="not\_experimental"  
 /product="dj1100H13.1.1 (KIAA1219 (similar to Drosophila  
 GH09358 and C. elegans D2085.5) isoform 1)"  
 /protein\_id="CAB41846.1"  
 /db\_xref="GI:4704208"  
 /translation="DSLESNISDSDSDNMDLPGILKQPSLTLELFPNHTDNLNSQ  
 RLSPPSRMRKLPQRPVPLPQTRVSVVVEYDDIEFPLSELMTETSTGCVETTFAN  
 SSTSLSRTTLEKEVPIFIHPLNTGLFRIKIQGATCKFNMVPLVDGCMVSRALGFL  
 VQTIVINICRRKLESDSYSPPHVRKQKIDIVNKNRQKLEPEFTTSLFQVGLKN  
 C55"  
 2753. .3187  
 /note="LIM4 repeat: matches 3867. .4299 of consensus"  
 4727. .5189  
 /gene="dj1100H13.1"  
 /note="match: GSS: Em:AAQ663797"  
 4732. .4875  
 /gene="dj1100H13.1"  
 /note="match: STS: Em:HSPF04E1"  
 4745. .5581  
 /gene="dj1100H13.1"  
 /note="match: GSS: Em:AAQ745415"  
 complement(4997. .5442)  
 /note="match: STS: Em:G33216"  
 complement(5210. .5903)  
 /note="match: GSS: Em:AAQ321022"  
 5322. .6464  
 /note="LIM4 repeat: matches 5388. .6322 of consensus"  
 6837. .6970  
 /note="MIR repeat: matches 32. .175 of consensus"  
 complement(6984. .7365)  
 /note="match: GSS: Em:AAQ807920"  
 7232. .7391  
 /note="LIM4 repeat: matches 7614. .7783 of consensus"  
 7610. .7684  
 /note="AluJb repeat: matches 39. .113 of consensus"  
 join(<8150. .8789,9265. .9431,10364. .>10484)  
 /gene="dj1100H13.1"  
 /product="dj1100H13.1.2 (KIAA1219, isoform 2)"  
 /note="match: ESTs: Em:AA536595"  
 /evidence="not\_experimental"  
 8567. .8616  
 /note="2 copies 25 mer 98% conserved"  
 12187. .12234  
 /note="12 copies 4 mer caca 89% conserved"  
 12536. .12571  
 /note="18 copies 2 mer aa 80% conserved"  
 12911. .12955



Query Match	Best Local Similarity	Mismatches	Score	DB 9;	Length
Matches	257;	Conservative	0;	Mismatches	29;
				Indels	2;
				Gaps	1;
Qy	13	GTCCGTGAGGCTGTGGCCCTGCAGCTGCGCCGCCATGGACGCTGCCGAGGTCGATTCCT	72		
Db	50001	GACTGTGAGCCCGAGCTGGCCACTGCGCCGCCATGGACGCTGCCGAGGTCGATTCCT	49942		
Qy	73	CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCGTGGACACAGATCTACCTCAT	132		
Db	49941	GGCTGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCGTGGACACAGATCTACCTCAT	49882		

Qy	133	CGGG--GGGGACCTGGGGCCCTTTTAAACCTCGTGTACCGTGGAAAGTGCCCTGTGGCTG	190
Db	49881	TGGGGTGGGTACCTGGGGCCCTTTTAAACCTCGTGTACCGTGGAAAGTGCCCTGTGGCTG	49822
Qy	191	GCATTAACTGAAACAAACAGACAAATCTCGCTCTCCCTCCAGAGTGGATGCATCTA	250
Db	49821	GCAATTAACTGAAACAAACAGACAAATCTCGCTCTCCCTCCAGAGTGGATGCATCTA	49762
Qy	251	GAAGAAGTGGAGAAGATGAGGATCATGAACGAAAGAAAGAACTTTT	298
Db	49761	GAAAGTGGAGAAGATGAGGATCATGAATGAAGAAAGAACTTTT	49714
RESULT 10			
AX336730		454 bp	DNA linear PAT 09-JAN-2002
LOCUS	Sequence 7239 from Patent WO0194629.		
DEFINITION	AX336730		
ACCESSION	AX336730.1	GI:18127449	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
JOURNAL	Horrigan, S., Soppet, D.R. and Weaver, Z.		
FEATURES	Cancer gene determination and therapeutic screening using signature		
source	gene sets		
	Patent: WO 0194629-A 7239 13-DEC-2001;		
	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
	1..454		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	128 a 113 c 116 g	97 t	
ORIGIN			
Query Match	74.5%;	Score 223.4;	DB 6; Length 454;
Best Local Similarity	99.6%;	Pred. No. 3e-48;	
Matches	224; Conservative	0; Mismatches	1; Indels
			0; Gaps
Qy	76	CGAGAAGGACCTGGTTACCATTTATCCCGAACCTTCAGTCTGGACAAGATCTACCTATCGG	135
Db	2	CGATAAGGACCTGGTTACCATTTATCCCGAACCTTCAGTCTGGACAAGATCTACCTATCGG	61
Qy	136	GGGGGACCTGGGGCCCTTTTAAACCTGTGTTTACCGTGGAAAGTGCCCTGTGGCTGGCGAT	195
Db	62	GGGGGACCTGGGGCCCTTTTAAACCTGTGTTTACCGTGGAAAGTGCCCTGTGGCTGGCGAT	121
Qy	196	TAACTGAAACAAAGACAGAAATGTGCCTGTCTCCCTCCAGAGTGGATGGATGTAGAAA	255
Db	122	TAACTGAAACAAAGACAGAAATGTGCCTGTCTCCCTCCAGAGTGGATGGATGTAGAAA	181
Qy	256	GTGGAGAAGTGAAGATCATGAACGAAAGGAAAGAACTTTTAC	300
Db	182	GTGGAGAAGTGAAGATCATGAACGAAAGGAAAGAACTTTTAC	226
RESULT 11			
AB097169		933 bp	DNA linear VRT 09-MAY-2003
LOCUS	AB097169		
DEFINITION	Xenopus laevis Psf2 gene, complete cds.		
ACCESSION	AB097169		
VERSION	AB097169.1	GI:29365480	
KEYWORDS			
SOURCE	Xenopus laevis (African clawed frog)		
ORGANISM	Xenopus laevis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
	Xenopodinae; Xenopus.		

REFERENCE  
AUTHORS Kubota,Y., Takase,Y., Komori,Y., Hashimoto,Y., Arata,T., Kamimura,Y., Araki,H. and Takisawa,H.  
TITLE A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication  
JOURNAL Genes Dev. 17 (9), 1141-1152 (2003)  
MEDLINE 22615695  
PUBMED 12730133  
REFERENCE  
AUTHORS Kubota,Y. and Takisawa,H.  
TITLE Direct Submission  
JOURNAL Submitted (28-NOV-2002) Yumiko Kubota, Osaka University, Department of Biology, Graduate School of Science; Machikaneyamacho 1-1, Toyonaka, Osaka 560-0043, Japan  
(E-mail:ykubota@bio.sci.osaka-u.ac.jp, Tel:81-6-6850-5554, Fax:81-6-6850-5554)

FEATURES  
source Location/Qualifiers  
1. .933  
/organism="Xenopus laevis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8355"  
21. .578  
/gene="Psf2"  
21. .578  
/gene="Psf2"  
/function="essential for DNA replication"  
/codon\_start=1  
/product="Psf2"  
/protein\_id="BAC66459.1"  
/db\_xref="GI:29365481"  
/translation="MDASEVEFLAEKEQVTVIPNFSLDKVLIGDLPFNPSLPVEV  
PLWLAINLQKQKRIIVPEWMDKLEAIRQERREFTFPMPSYMYELKLLLNH  
AADNIPKADEITLVKDTWDTRIAKLRISADSFVKQEAHAKLDNLTLWEINTIGTFE  
TESLNHYKLRTSLONPEBGSQDY"  
BASE COUNT 282 a 200 c 208 g 243 t  
ORIGIN

Query Match 55.8%; Score 167.4; DB 5; Length 933;  
Best Local Similarity 78.2%; Pred. No. 1.8e-33;  
Matches 201; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 44 GCATGAGCGCTGCCAGAGTCGAATTCCTCGCGGAGAGGAGCTGGTTACCATATCCCC 103  
Db 18 GTACATGATGCGCTCTGAGTGAGTTCTTGGGTGAGAGGAGCAAGTACCGTCATCCCC 77  
QY 104 AACTTCAGTCTGCACAAATCATACCTCATCGGGGGGAGCTGGGGCTTTTAAACCTGGT 163  
Db 78 AACTTAGTCTGGACAAGATCTATCTCATTTGGGGGTGACTTGGGCCCTTTAATCCTAGT 137  
QY 164 TTACCGCTGGAAGTGGCCCTGTGGCTGGCGGATTAACCTGAAACAAAGACAGAAATGTCG 223  
Db 138 TTACCTGTTGAGGTTCTCTGTGGTTGGCCATCACTGAACAAAGGAGAAATGTCGA 197  
QY 224 CTGCTCCCTCCAGAGTGGATGATGATGAGAAAGTTGGAGAAGATGAGGATCATGAACGA 283  
Db 198 ATFTGCTCCCTCCAGAGTGGATGATGATGAGAAACTAGAGGCTATCCGAGACCAAGAACGC 257  
QY 284 AAGGAGAAACTTTTAC 300  
Db 258 AGAGGAGACGCTTAC 274

RESULT 12  
AC009068/c  
LOCUS AC009068 160356 bp DNA linear PRI 27-APR-2001  
DEFINITION Homo sapiens chromosome 16 clone RP11-314K3, complete sequence.  
ACCESSION AC009068  
VERSION AC009068.10 GI:13811892  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (27-APR-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Apr 27, 2001 this sequence version replaced gi:12000278.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.3.  
STS Content:  
WI-11796 G24048  
SHGC-37143 G30481.

FEATURES  
source Location/Qualifiers  
1. .160356  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-314K3"  
BASE COUNT 37469 a 43134 c 43308 g 36445 t  
ORIGIN

Query Match 41.7%; Score 125.2; DB 9; Length 160356;  
Best Local Similarity 94.2%; Pred. No. 3.7e-22;  
Matches 130; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 13 GTCGGTGAGGCTCTGCGCTGCAGCTCGCGCGGCATGGAGCTGCCGAGTGCGAATTCCT 72  
Db 60646 GACCTGTGAGGCTCTGCGCTGCAGCTCGCGCGGCATGGAGCTGCCGAGTGCGAATTCCT 60587  
QY 73 GCCCGAGAAGAGCTGGTTACCATTTACCATTTACCATTTACCATTTACCATTTACCTCAT 132  
Db 60586 CGCCGAGAAGAGCTGGTTACCATTTACCATTTACCATTTACCATTTACCATTTACCTCAT 60527  
QY 133 CGGGGGGAGCTGGGGGCC 150  
Db 60526 CGGGGTGAGGCGCTGGGCC 60509

RESULT 13  
AC123908/c  
LOCUS AC123908 185321 bp DNA linear PRI 28-MAR-2003  
DEFINITION Homo sapiens chromosome 16 clone CTD-2542L18, complete sequence.  
ACCESSION AC123908  
VERSION AC123908.3 GI:29336199  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUN-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
AUTHORS DOE Joint Genome Institute.

AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (06-SEP-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 185321)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Mar 28, 2003 this sequence version replaced gi:22748376.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos  
National Laboratory  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.  
FEATURES  
source Location/Qualifiers  
1..185321  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="CTD-2542L18"  
BASE COUNT 40674 a 50904 c 51869 g 41874 t  
ORIGIN  
Query Match 41.7%; Score 125.2; DB 9; Length 185321;  
Best Local Similarity 94.2%; Pred. No. 3.7e-22;  
Matches 130; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 13 CTCGGTGAGGCTCTGGCTGAGCTGCGCGCCATGCGAGCTGCCGAGGTCGAATTCCT 72  
Db 105811 GACCGTGAGGCTCTGGCTGAGCTGCGCGCCATGCGAGCTGCCGAGGTCGAATTCCT 105752  
QY 73 CGCCGAGAGGAGCTGGTTACCATATCCCACTTCCAGTCTGACAGATCTACCTCAT 132  
Db 105751 CGCCGAGAGGAGCTGGTTACCATATCCCACTTCCAGTCTGACAGATCTACCTCAT 105692  
QY 133 CGGGGGGAGCCTGGGGCC 150  
Db 105691 CGGGGTGAGGCTGGGCC 105674  
RESULT 14  
AC137771/c 245210 bp DNA linear HTG 03-DEC-2002  
LOCUS  
DEFINITION Homo sapiens chromosome 16 clone CTD-2139A24, WORKING DRAFT  
SEQUENCE, 45 unordered pieces.  
ACCESSION AC137771  
VERSION AC137771.1 GI:26006513  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 245210)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 245210)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-DEC-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT -----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 680949

Center clone name: CITB-HI\_2139A24  
-----  
Summary Statistics  
Consensus quality: 221831 bases at least Q40  
Consensus quality: 226530 bases at least Q30  
Consensus quality: 229658 bases at least Q20  
Estimated insert size: 170000; agarose-fp estimation  
Quality coverage: 12.32 in Q20 bases; agarose-fp estimation  
Quality coverage: 12.32 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 45 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1084: contig of 1084 bp in length  
\* 1085 1184: gap of unknown length  
\* 1185 2248: contig of 1064 bp in length  
\* 2249 2348: gap of unknown length  
\* 2349 3435: contig of 1087 bp in length  
\* 3436 3535: gap of unknown length  
\* 3536 4637: contig of 1102 bp in length  
\* 4638 4737: gap of unknown length  
\* 4738 5816: contig of 1079 bp in length  
\* 5817 5916: gap of unknown length  
\* 5917 6951: contig of 1035 bp in length  
\* 6952 7051: gap of unknown length  
\* 7051 8599: contig of 1548 bp in length  
\* 8600 8699: gap of unknown length  
\* 8700 10508: contig of 1809 bp in length  
\* 10509 10608: gap of unknown length  
\* 10609 11738: contig of 1130 bp in length  
\* 11739 11838: gap of unknown length  
\* 11839 13754: contig of 1916 bp in length  
\* 13755 13854: gap of unknown length  
\* 13855 14996: contig of 1142 bp in length  
\* 14997 15096: gap of unknown length  
\* 15097 16852: contig of 1756 bp in length  
\* 16853 16952: gap of unknown length  
\* 16953 18502: contig of 1550 bp in length  
\* 18503 18602: gap of unknown length  
\* 18603 20737: contig of 2134 bp in length  
\* 20737 20836: gap of unknown length  
\* 20837 22243: contig of 1407 bp in length  
\* 22244 22343: gap of unknown length  
\* 22344 24310: contig of 1967 bp in length  
\* 24311 24410: gap of unknown length  
\* 24411 26675: contig of 2265 bp in length  
\* 26676 26775: gap of unknown length  
\* 26776 28257: contig of 1482 bp in length  
\* 28258 28357: gap of unknown length  
\* 28358 29668: contig of 1311 bp in length  
\* 29669 29768: gap of unknown length  
\* 29769 31503: contig of 1735 bp in length  
\* 31504 33962: contig of 2359 bp in length  
\* 33963 34062: gap of unknown length  
\* 34063 36099: contig of 2037 bp in length  
\* 36100 36199: gap of unknown length  
\* 36200 38832: contig of 2633 bp in length  
\* 38833 38932: gap of unknown length  
\* 38933 41464: contig of 2532 bp in length  
\* 41465 41564: gap of unknown length  
\* 41565 44085: contig of 2521 bp in length  
\* 44086 44185: gap of unknown length  
\* 44186 47028: contig of 2843 bp in length  
\* 47029 47128: gap of unknown length  
\* 47129 49395: contig of 2267 bp in length  
\* 49396 49495: gap of unknown length  
\* 49496 51435: contig of 1940 bp in length  
\* 51435



may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWGE

Center clone name: CH230-247G13

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 198554 bases at least Q40

Consensus quality: 206661 bases at least Q30

Consensus quality: 201900 bases at least Q20

Estimated insert size: 210626; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 1 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 207307: contig of 207307 bp in length.

#### FEATURES

source

1. 207307

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-247G13"

87039..177061

/note="clone\_boundary"

clone\_end:Sp6

site:Mbol

end\_sequence:Bzl68551"

150202..150357

/note="clone\_boundary"

clone\_end:T7

site:Mbol

end\_sequence:Bzl68550"

BASE COUNT 43941 a 54080 c 55681 g 48886 t 4719 others

ORIGIN

Query Match 32.3%; Score 97; DB 2; Length 207307;

Best Local Similarity 87.6%; Pred. No. 1e-14; Indels 0; Gaps 0;

Matches 106; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 136 GGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGTGAAGTGCCTGTGGCTGGCGAT 195

Db 205359 GGGGGACCTGGGGCCCTTCAACCCCGCTTACCCGTGGACGTGCCCTGTGGCTGGCCAT 205300

QY 196 TAACTGAAACAAAGACAGAAATGTCGCTGTCTCCCTCCAGAGTGGATGGATGTAGAAAA 255

Db 205299 TAACTGAAACAGACAGACAGAGTGGCGTCTGCTACTCTCCAGAGTGGATGGATGGGTGA 205240

QY 256 G 256

Db 205239 G 205239

Search completed: October 14, 2003, 17:02:29  
Job time : 1481 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:05:16 ; Search time 36.6667 Seconds  
(without alignments)  
3611.317 Million cell updates/sec

Title: US-09-854-124-5  
Perfect score: 300  
Sequence: 1 acgaataccgagccctggtc.....tgaggtactctggttctgg 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	11.7	2241	2	US-08-838-219B-20
2	35	11.7	2241	3	US-09-233-336A-20
3	35	11.7	2241	3	US-09-233-752A-20
4	35	11.7	2241	3	US-09-402-036-20
5	35	11.7	2241	4	US-09-904-226-20
6	35	11.7	2370	2	US-08-838-219B-19
7	35	11.7	2370	3	US-09-233-336A-19
8	35	11.7	2370	3	US-09-233-752A-19
9	35	11.7	2370	3	US-09-402-036-19
10	35	11.7	2370	4	US-09-904-226-19
11	35	11.7	2403	1	US-08-471-033-30
12	35	11.7	2403	1	US-08-471-044-30
13	35	11.7	2403	2	US-08-463-483A-30
14	35	11.7	2403	2	US-08-471-046A-30
15	35	11.7	2403	2	US-08-470-566B-30
16	35	11.7	2403	2	US-08-838-219B-7
17	35	11.7	2403	2	US-08-469-334-30
18	35	11.7	2403	3	US-09-300-529-30
19	35	11.7	2403	3	US-09-233-336A-7
20	35	11.7	2403	3	US-09-233-752A-7
21	35	11.7	2403	3	US-09-402-036-7
22	35	11.7	2403	4	US-09-904-226-7
23	31.4	10.5	1296	4	US-09-252-991A-7149
24	31.4	10.5	1419	4	US-09-252-991A-7347
25	31.4	10.5	2118	4	US-09-016-434-1304
26	31.4	10.5	2748	4	US-09-252-991A-7208
27	31.4	10.5	2831	2	US-08-906-713-1

Sequence 43, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 33, Appli  
Sequence 3, Appli  
Sequence 5, Appli  
Sequence 31, Appli  
Sequence 21, Appli  
Sequence 2, Appli  
Sequence 1192, Ap  
Patent No. 5472691  
Sequence 1, Appli  
Patent No. 5472691  
Sequence 12012, A  
Sequence 11975, A  
Sequence 11906, A  
Patent No. 5472691  
Sequence 20, Appli

ALIGNMENTS

RESULT 1  
US-08-838-219B-20  
; Sequence 20, Application US/08838219B  
; Patent No. 5877012  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Gregory W  
; APPLICANT: Koziel, Michael G  
; APPLICANT: Mullins, Martha A  
; APPLICANT: Nye, Gordon J  
; APPLICANT: Carr, Brian  
; APPLICANT: Desai, Nalini M  
; APPLICANT: Kostichka, N. Kristy  
; APPLICANT: Duck, Nicholas B  
; APPLICANT: Estruch, Juan J  
; TITLE OF INVENTION: A No. 5877012el Class of Proteins for the  
; CONTROL OF PLANT PESTS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,219B  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/463,483  
; FILING DATE: 06-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/314,594  
; FILING DATE: 09-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/218,018  
; FILING DATE: 23-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/037,057  
; FILING DATE: 25-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Gary M.  
; REGISTRATION NUMBER: 40,403  
; REFERENCE/DOCKET NUMBER: CGC 1925  
; TELECOMMUNICATION INFORMATION:

28 31 10.3 717 4 US-09-513-783A-43  
29 31 10.3 1389 2 US-08-023-980B-3  
30 31 10.3 1389 2 US-08-486-953A-3  
31 31 10.3 1623 4 US-09-513-783A-33  
32 31 10.3 2439 4 US-09-513-783A-3  
33 31 10.3 2439 4 US-09-513-783A-5  
34 31 10.3 3171 4 US-09-513-783A-21  
35 31 10.3 4833 4 US-09-513-783A-21  
36 30.8 10.3 1625 2 US-08-709-923-2  
37 30.8 10.3 2889 4 US-09-016-434-1192  
38 29.8 9.9 666 6 5472691-4  
39 29.8 9.9 569 3 US-08-556-965-1  
40 29.8 9.9 723 6 5472691-7  
41 29.8 9.9 792 4 US-09-252-991A-12012  
42 29.8 9.9 1017 4 US-09-252-991A-11975  
43 29.8 9.9 1104 4 US-09-252-991A-11906  
44 29.8 9.9 1396 6 5472691-1  
45 29.8 9.9 10079 2 US-08-476-866-20





```

RESULT 5
US-09-904-226-20
; Sequence 20, Application US/09904226
; Patent No. 6429360
; GENERAL INFORMATION:
; APPLICANT: Estruch, Juan J.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Desai, Nalini
; APPLICANT: Koziel, Michael
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: Plant Pest Control
; FILE REFERENCE: S-21284D
; CURRENT APPLICATION NUMBER: US/09/904,226
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/EP98/01952
; PRIOR FILING DATE: 1998-04-02

```

Query Match	11.7%	Score 35;	DB 4;	Length 2241;
Best Local Similarity	59.6%	Pred. No. 0.14;		
Matches	59;	Conservative 0;	Mismatches 40;	Indels 0; Gaps 0;
QY	77	GAGACAGCAGGAGGCACATGCCAAGCTGGGATACCTTGACCTTGATGGAGATCAACACG	136	
Db	1950	GAGCCAGAGCGGCGAGGCGCTGGGCGCAACTTCATCATCTCTGGAGATCAGCCGAG	2009	
QY	137	CGGGAGCTTTCCTCACACAAAGCGCTCAACCAACATGTACAA	175	
Db	2010	CGAGAGCTCTGAGCCCGGAGCTGATCAACACCAACA	2048	

Db, 2010 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCCAA 2048

Db 1950 GAGCCAGAACGGCGACGAGGCCTGGGGCGACAATTATCATCTCGGAGATCAGCCCCGAG 2009



```
; APPLICANT: Koziel, Michael
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: Plant Pest Control
; FILE REFERENCE: S-21284D
; CURRENT APPLICATION NUMBER: US/09/904,226
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/EP98/01952
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 08/838,219
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,263
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,265
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/463,483
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/314,594
; PRIOR FILING DATE: 1994-09-09
; PRIOR APPLICATION NUMBER: 08/218,018
; PRIOR FILING DATE: 1994-03-24
; PRIOR APPLICATION NUMBER: 08/037,057
; PRIOR FILING DATE: 1993-03-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
; OTHER INFORMATION: encoding VIP3A(b)
US-09-904-226-19
```

```
Query Match 11.7%; Score 35; DB 4; Length 2370;
Best Local Similarity 59.6%; Pred. No. 0.15;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 77 GAGACAGCAGGAGCCATGCGCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACGAG 136
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1950 GAGCCAGAACGGCGACGAGCGCTGGGGCGACAACTTCATCATCTCTGGAGATCAGCCCGAG 2009
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 CGGCACTTCTCTACACAAAGCGCTCAACCAACATGTACAA 175
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2010 CGAGAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2048
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-08-471-033-30
; Sequence 30, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA
; OTHER INFORMATION: sequence encoding VIP3A(a)"
US-08-471-033-30

Query Match 11.7%; Score 35; DB 1; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.15;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 77 GAGACAGCAGGAGCCATGCGCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACGAG 136
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1966 GAGCCAGAACGGCGACGAGCGCTGGGGCGACAACTTCATCATCTCTGGAGATCAGCCCGAG 2025
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 CGGCACTTCTCTACACAAAGCGCTCAACCAACATGTACAA 175
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2026 CGAGAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-471-044-30
; Sequence 30, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,044
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA
; OTHER INFORMATION: sequence encoding VIP3A(a)"
;
US-08-471-044-30

Query Match 11.7%; Score 35; DB 2; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.15;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 77 GAGACAGCAGGAGCAGCATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACCAG 136
Db 1966 GAGCCAGAACGGCGACGAGGCGCTGGGGCGACAACTTCATCATCTCTGGAGATCAGCCCGAG 2025

QY 137 CGGACTTTCTCTACACAAGCGCTCAACACCATGTACAA 175
Db 2026 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064

RESULT 13
US-08-463-483A-30
; Sequence 30, Application US/08463483A
; Patent No. 5849870
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Kozziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,483A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA
; OTHER INFORMATION: sequence encoding VIP3A(a)"
;
US-08-463-483A-30

Query Match 11.7%; Score 35; DB 2; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.15;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 77 GAGACAGCAGGAGCAGCATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACCAG 136
Db 1966 GAGCCAGAACGGCGACGAGGCGCTGGGGCGACAACTTCATCATCTCTGGAGATCAGCCCGAG 2025

QY 137 CGGACTTTCTCTACACAAGCGCTCAACACCATGTACAA 175
Db 2026 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064

RESULT 14
US-08-471-046A-30
; Sequence 30, Application US/08471046A
; Patent No. 5866326
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Kozziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
```



Db 2026 CGAGAGCTGTGAGCCCGGAGCTGATCAACACCAACAA 2064

Search completed: October 14, 2003, 17:04:21  
Job time : 38.6667 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 14:17:11 ; Search time 143.333 Seconds  
(without alignments)  
5431.000 Million cell updates/sec

Title: US-09-854-124-5

Perfect score: 300

Sequence: 1 acgaataccgagccctggtc.....tgaggtaactcgtggtctgg 300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	9	US-09-854-124-5
2	300	100.0	1210	9	US-09-925-701-214
3	239	79.7	484	11	US-09-918-995-19938
4	214	71.5	200	9	US-09-854-124-7
5	199	66.3	382	10	US-09-796-692-9460
6	199	66.3	382	14	US-10-040-862-9460
7	194	64.7	597	10	US-09-796-692-9460
8	194	64.7	597	14	US-10-040-862-9460
9	66	22.0	362	10	US-09-920-300A-682
10	66	22.0	362	10	US-09-998-598-2284
11	66	22.0	362	12	US-10-099-946-682
12	66	22.0	362	13	US-10-033-528-682
13	39.2	13.1	820	13	US-10-027-632-151845
14	39.2	13.1	820	13	US-10-027-632-151846
15	32.8	10.9	468	9	US-09-854-761-789
16	32.8	10.9	737	9	US-09-854-761-17572

c 17	32.4	10.8	1188	14	US-10-156-761-5491	Sequence 5491, Ap
c 18	32.4	10.8	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 19	31.6	10.5	650	13	US-10-027-632-246523	Sequence 246523,
c 20	31.6	10.5	831	14	US-10-156-761-4064	Sequence 4064, Ap
c 21	31.4	10.5	360	10	US-09-920-455-112	Sequence 112, App
c 22	31.4	10.5	517	10	US-09-998-598-1387	Sequence 1387, Ap
c 23	31.4	10.5	596	13	US-10-027-632-50355	Sequence 50355, A
c 24	31.4	10.5	596	13	US-10-027-632-50356	Sequence 50356, A
c 25	31.4	10.5	1722	10	US-09-912-672A-3	Sequence 3, Appli
c 26	31.4	10.5	2478	12	US-10-063-735-163	Sequence 163, App
c 27	31.4	10.5	2478	12	US-10-063-526-163	Sequence 163, App
c 28	31.4	10.5	2478	12	US-10-063-586-163	Sequence 163, App
c 29	31.4	10.5	2478	12	US-10-063-510-163	Sequence 163, App
c 30	31.4	10.5	2478	12	US-10-063-514-163	Sequence 163, App
c 31	31.4	10.5	2478	12	US-10-063-516-163	Sequence 163, App
c 32	31.4	10.5	2478	12	US-10-063-523-163	Sequence 163, App
c 33	31.4	10.5	2478	12	US-10-063-527-163	Sequence 163, App
c 34	31.4	10.5	2478	12	US-10-063-528-163	Sequence 163, App
c 35	31.4	10.5	2478	12	US-10-063-529-163	Sequence 163, App
c 36	31.4	10.5	2478	12	US-10-063-536-163	Sequence 163, App
c 37	31.4	10.5	2478	12	US-10-063-540-163	Sequence 163, App
c 38	31.4	10.5	2478	12	US-10-063-546-163	Sequence 163, App
c 39	31.4	10.5	2478	12	US-10-063-562-163	Sequence 163, App
c 40	31.4	10.5	2478	12	US-10-063-564-163	Sequence 163, App
c 41	31.4	10.5	2478	12	US-10-063-565-163	Sequence 163, App
c 42	31.4	10.5	2478	12	US-10-063-568-163	Sequence 163, App
c 43	31.4	10.5	2478	12	US-10-063-570-163	Sequence 163, App
c 44	31.4	10.5	2478	12	US-10-063-577-163	Sequence 163, App
c 45	31.4	10.5	2478	12	US-10-063-577-163	Sequence 163, App

#### ALIGNMENTS

#### RESULT 1

US-09-854-124-5  
; Sequence 5, Application US/09854124  
; Patent No. US20020076735A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Innis, Michael A.  
; APPLICANT: Garcia, Pablo Dominquez  
; APPLICANT: Sudduth-Klinger, Julie  
; APPLICANT: Reinhard, Christoph  
; APPLICANT: Giese, Klaus  
; APPLICANT: Randazzo, Filippo  
; APPLICANT: Kennedy, Giulia C.  
; APPLICANT: Pot, David  
; APPLICANT: Kassam, Altaf  
; APPLICANT: Lamsom, George  
; APPLICANT: Drmanac, Radoje  
; APPLICANT: Crkvenjakov, Radomir  
; APPLICANT: Dickson, Mark  
; APPLICANT: Drmanac, Snezana  
; APPLICANT: Labat, Ivan  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Kita, David  
; APPLICANT: Garcia, Veronica  
; APPLICANT: Jones, Lee William  
; APPLICANT: Stache-Crain, Birgit  
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods Using  
; FILE REFERENCE: 2300-1490  
; CURRENT APPLICATION NUMBER: US/09/854,124  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 09/400,947  
; PRIOR FILING DATE: 1999-09-22  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 300  
; TYPE: DNA

```

; ORGANISM: Homo sapiens
US-09-854-124-5

Query Match      100.0%; Score 300; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-90;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  ACGAAATCCGGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60
Db      1  ACGAAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60

QY      61  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGACCTTGA 120
Db      61  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGACCTTGA 120

QY      121  TGGAGATCAACACACAGCGGGGACTTCTTCACAAAGCGCTCAACACATGTACAAACTCC 180
Db      121  TGGAGATCAACACACAGCGGGGACTTCTTCACAAAGCGCTCAACACATGTACAAACTCC 180

QY      181  GCAGCAACTCCAGCCCTCTGGAGAGTACTCAAGTCTCAGGACTTCTAGAGAAAGGCGCTGCT 240
Db      181  GCAGCAACTCCAGCCCTCTGGAGAGTACTCAAGTCTCAGGACTTCTAGAGAAAGGCGCTGCT 240

QY      241  GCAGGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATCAGGTACTCGTGGTTCTGG 300
Db      241  GCAGGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATCAGGTACTCGTGGTTCTGG 300

RESULT 2
US-09-925-301-214
; Sequence 214, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 214
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-214

```

```

RESULT 3
US-09-918-995-19938
; Sequence 19938, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19938
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-19938

Query Match          79.7%; Score 239; DB 11; Length 484;
Best Local Similarity 100.0%; Pred No. 4.1e-70;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60
      |||
Db      245  ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 304

Qy      61  CTGCTGACAGCTTTGTGAGACACAGCAGGAGGCACATGCCAAGCTGGTAACCTTGACCTTGA 120
      |||
Db      305  CTGCTGACAGCTTTGTGAGACACAGCAGGAGGCACATGCCAAGCTGGTAACCTTGACCTTGA 364

Qy      121  TGGAGATCAACACCCAGCGGGACTTTCTCTCACAAAGCGCTCAACACACATGTACAAACTCC 180
      |||
Db      365  TGGAGATCAACACCCAGCGGGACTTTCTCTCACAAAGCGCTCAACACACATGTACAAACTCC 424

Qy      181  GCACGAACCTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGG 239
      |||
Db      425  GCACGAACCTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGG 483

RESULT 4
US-09-854-124-7
; Sequence 7, Application US/099854124
; Patent No. US20020076735A1

```

Query Match	100.0%;	Score 300;	DB 9;	Length 1210;
Best Local	Similarity 100.0%;	Pred. No. 1.8e-90;		
Matches 300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT	60	
Db	412	ACGAAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT	471	
QY	61	CTGCTGACAGCTTTGTGTAGACACGAGAGGACACATGCCAAGCTGGATAACTTGGACCTTGA	120	
Db	472	CTGCTGACAGCTTTGTGTAGACACGAGAGGACACATGCCAAGCTGGATAACTTGGACCTTGA	531	
QY	121	TGGAGATCAACACCCAGCGGGACTTTCCTCACAAAGCGCTCAAGCACATGTACAAACTCC	180	
Db	532	TGGAGATCAACACCCAGCGGGACTTTCCTCACAAAGCGCTCAAGCACATGTACAAACTCC	591	
QY	181	GCACGAACCTCCAGCCTCTCGGAGGATCTCAGTCTCAGGACTTCTTAGAGAAAGGCCCTGGT	240	
Db	592	GCACGAACCTCCAGCCTCTCGGAGGATCTCAGTCTCAGGACTTCTTAGAGAAAGGCCCTGGT	651	
QY	241	GCAGGGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTCG	300	
Db	652	GCAGGGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTCG	711	

```
; FILE REFERENCE: 2300-1490
; CURRENT APPLICATION NUMBER: US/09/854,124
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/400,947
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-124-7

Query Match      71.5%; Score 214.4; DB 9; Length 300;
Best Local Similarity 94.6%; Pred. No. 6.4e-62;
Matches 244; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ACGAATCCGACCGCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60
   |||||
Db 31 ACGAATCCGACCGCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 90
   |||||
QY 61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCGACATGCCAAAGCTGGATAACTTTGACCTTGA 120
   |||||
Db 91 CTGCTGACAGCTTTGTGACACAGCAGGAGGCGACATGCCAAAGCTGGATAACTTTGACCTTGA 150
   |||||
QY 121 TGGAGATCAACACGACGGGACTTTCTCTCACACAGGCGCTCAACCCACATGTACAAACTCC 180
   |||||
Db 151 TGGAGATCAACACGACGGGACTTTCTCTCACACAGGCGCTCAACCCACATGTACAAACTCC 210
   |||||
QY 181 GCAGGAACCTCCAGCCTCTGGAGAGTACTCAGCTCAGGACTCTAGAGAAAGGCGCTGT 240
   |||||
Db 211 GCAGGAACCTCCAGCCTCTGGAGAGTACTCAGCTCAGGACTCTAGAGAAAGGCGCTGT 267
   |||||

QY 241 GCAGGCGGCTGTGCTGGGG 258
   |||||
Db 268 GCAGGCGCTGTGCTGGGG 285

RESULT 5
US-09-796-692-9460/c
; Sequence 9460, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04

; FILE REFERENCE: 2300-1490
; CURRENT APPLICATION NUMBER: US/09/854,124
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/400,947
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9460
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (72)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (80)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (89)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (95)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (122)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (129)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (131)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (155)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (173)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (222)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (235)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (236)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (237)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (245)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (251)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (305)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (312)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (321)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (334)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
```

; LOCATION: (352)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (354)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (373)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (375)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-9460

Query Match 66.3%; Score 199; DB 10; Length 382;  
Best Local Similarity 94.4%; Pred. No. 1e-56;  
Matches 221; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
  
QY 3 GAATCCGGACCCCTGGTCAAGGATATGCGACACTCGTATAGCCAACTCCGAGTGCT 62  
|||||  
DB 234 GAATCCGGACCCCTGGTCAAGGATATGCGACACTCGTATAGCCAACTCCGAGTGCT 175  
|||||  
QY 63 GCTGACAGCTTTGTGAGACA-GCAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGAT 121  
|||||  
DB 174 GNTGACAGCTTTGTGAGACNCGCAGGAGGCACATGCCAAGCTGNANAACTTGNCTTGAT 115  
|||||  
QY 122 GGAGATCAACACAGCGGGAGCTTCTCTCACACA-AGCGCTCAACACCATGTACAAACTCC 180  
|||||  
DB 114 GGAGATCAACACAGCGGGAGCTTCTCTCACACATNGCGCTCANGCATGTACAAACTCC 55  
|||||  
QY 181 GCAGCAACTCCAGCCCTCTGGAGAGTACTCTAGTCTCAGGACTTCTAGAGAAAG 234  
|||||  
DB 54 GCAGCAACTCCAGCCCTCTGGAGAGTACTCTAGTCTCAGGACTTCTAGAGAAAG 1  
|||||

## RESULT 6

US-10-040-862-9460/c  
; Sequence 9460, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040.862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07

; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9460  
; LENGTH: 382  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (11)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (72)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (80)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (89)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (95)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (122)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (129)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (131)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (155)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (173)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (222)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (235)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (236)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (237)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (245)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (251)  
; OTHER INFORMATION: n=A,T,C or G  
; FEATURE:

NAME/KEY: unsure  
LOCATION: (261)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (305)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (312)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (321)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (334)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (352)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (354)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (373)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (375)  
OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-9460

Query Match 66.3%; Score 199; DB 14; Length 382;

Best Local Similarity 94.4%; Pred. No. 1e-56;

Matches 221; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 3 GAAATCCGACCTGGTCAAGCATATGTGGGACACTCGTATAGCCAAACTCCGAGTGCT 62  
DB 234 GAAATCCGACCTGGTCAAGCATATGTGGGACACTCGTATAGCCAAACTCCGAGTGCT 175  
QY 63 GGTGACAGCTTTGTGAGACA-GCAGGAGGCACATGCCAAGCTGGGATACTTGACCTTGAT 121  
DB 174 GGTGACAGCTTTGTGAGACNCGAGGAGGCACATGCCAAGCTGNANAAGCTTGNCTTGAT 115  
QY 122 GGAGATCAACACCGAGGGGACTTTCTCACACA-AGCGCTCAACACCATGTACAAACTCC 180  
DB 114 GGAGATCAACACCGAGGGGCTTTCTCACATNGCGCTCANCCACATGTACAAACTCC 55  
QY 181 GCACGAACCTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGG 234  
DB 54 GCACGAACCTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGG 1

RESULT 7

US-09-796-692-4407

Sequence 4407, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4407  
LENGTH: 597  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (583)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (592)  
OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-4407

Query Match 64.7%; Score 194; DB 10; Length 597;

Best Local Similarity 99.0%; Pred. No. 5.7e-55;

Matches 194; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACGAAATCCGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGCT 60  
DB 402 ACGAAATCCGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGCT 461  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGGATTAACCTTGA 120  
DB 462 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGGATTAACCTTGA 521  
QY 121 TGGAGATCAACACCGAGGGGACTTTCTCACACAAGCGCTCAACCATGTACAAACTCC 180  
DB 522 TGGAGATCAACACCGAGGGGACTTTCTCACACAAGCGCTCAACCATGTACAAACTCC 581  
QY 181 GCACGAACCTCCAGCC 196  
DB 582 GNACGAACCTNCAGCC 597

RESULT 8

US-10-040-862-4407

Sequence 4407, Application US/10040862  
Publication No. US20030078396A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther  
FILE REFERENCE: 014058-01352085  
CURRENT APPLICATION NUMBER: US/10/040,862  
CURRENT FILING DATE: 2001-11-06

;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 4407  
;; LENGTH: 597  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (583)  
;; OTHER INFORMATION: n=A,T,C or G  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (592)  
;; OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-4407

Query Match 64.7%; Score 194; DB 14; Length 597;  
Best Local Similarity 99.0%; Pred. No. 5.7e-55;  
Matches 194; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
Db 402 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 461  
Qy 61 CTGCTGACAGCTTTGTGACAGCAGCAGGACACATGCCAAGCTGGATAACTTGACCTTGA 120  
Db 462 CTGCTGACAGCTTTGTGACAGCAGCAGGACACATGCCAAGCTGGATAACTTGACCTTGA 521  
Qy 121 TGGAGATCAACACCCAGCGGGACTTCTTCACACAAGCGCTCAACCATATGACAACTCC 180  
Db 522 TGGAGATCAACACCCAGCGGGACTTCTTCACACAAGCGCTCAACCATATGACAACTCC 180  
Qy 181 GCACGAACCTCCAGCC 196  
Db 582 GNACGAACCTCCAGCC 597

RESULT 9  
US-09-920-300A-682/c  
; Sequence 682, Application US/09920300A  
; Patent No. US20020136728A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secretist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

;; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
;; FILE REFERENCE: 210121.547  
;; CURRENT APPLICATION NUMBER: US/09/920,300A  
;; CURRENT FILING DATE: 2001-07-31  
;; NUMBER OF SEQ ID NOS: 1789  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 682  
;; LENGTH: 362  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-920-300A-682

Query Match 22.0%; Score 66; DB 10; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.6e-12;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 235 CCTGGTCAGCGCGCTTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTCGTGG 294  
Db 362 CCTGGTCAGCGCGCTTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTCGTGG 303  
Qy 295 TTCTGG 300  
Db 302 TTCTGG 297

RESULT 10  
US-09-998-598-2284/c  
; Sequence 2284, Application US/09998598  
; Patent No. US20020150922A1  
; GENERAL INFORMATION:  
; APPLICANT: Stolk, John A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Meagher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.561  
; CURRENT APPLICATION NUMBER: US/09/998,598  
; CURRENT FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 2606  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 2284  
; LENGTH: 362  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-998-598-2284

Query Match 22.0%; Score 66; DB 10; Length 362;  
Best Local Similarity 100.0%; Pred. No. 4.6e-12;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 235 CCTGGTCAGCGCGCTTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTCGTGG 294  
Db 362 CCTGGTCAGCGCGCTTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTCGTGG 303  
Qy 295 TTCTGG 300  
Db 302 TTCTGG 297

RESULT 11  
US-10-099-926-682/c  
; Sequence 682, Application US/10099926  
; Publication No. US2003016064A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Gordon E.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Secretist, Heather  
; APPLICANT: Jiang, Yugu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.547C2

```
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 682
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-682

Query Match          22.0%; Score 66; DB 12; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CCTGGTCAGCGCGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 294
Db 362 CCTGGTCAGCGCGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 303

QY 295 TTCTGG 300
Db 302 TTCTGG 297

RESULT 12
US-10-033-528-682/c
; Sequence 682, Application US/10033528
; Publication No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 682
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-682

Query Match          22.0%; Score 66; DB 13; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CCTGGTCAGCGCGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 294
Db 362 CCTGGTCAGCGCGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 303

QY 295 TTCTGG 300
Db 302 TTCTGG 297

RESULT 13
US-10-027-632-151845/c
; Sequence 151845, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151845
; LENGTH: 820
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151845

Query Match          13.1%; Score 39.2; DB 13; Length 820;
Best Local Similarity 60.2%; Pred. No. 0.006;
Matches 65; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 23 GGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACA 82
Db 180 GAAGAAGTGTGAGCGCTGTTTCACCCAAAGTCTGGTGAAGGGCCCCCTACTTGTCTGCA 121

QY 83 GCAGGAGGCACATGCCAAGCTTGATAACTTGACCTTGATGGAGATCAA 130
Db 120 GGACGATGCCAGGCCAAGGTGGTCAACATCACCATCGTGCGCCCTCAA 73
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151845
; LENGTH: 820
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151845

Query Match          13.1%; Score 39.2; DB 13; Length 820;
Best Local Similarity 60.2%; Pred. No. 0.006;
Matches 65; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 23 GGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACA 82
Db 180 GAAGAAGTGTGAGCGCTGTTTCACCCAAAGTCTGGTGAAGGGCCCCCTACTTGTCTGCA 121

QY 83 GCAGGAGGCACATGCCAAGCTTGATAACTTGACCTTGATGGAGATCAA 130
Db 120 GGACGATGCCAGGCCAAGGTGGTCAACATCACCATCGTGCGCCCTCAA 73

RESULT 14
US-10-027-632-151846/c
; Sequence 151846, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151846
; LENGTH: 820
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151846

Query Match          13.1%; Score 39.2; DB 13; Length 820;
Best Local Similarity 60.2%; Pred. No. 0.006;
Matches 65; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 23 GGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACA 82
Db 180 GAAGAAGTGTGAGCGCTGTTTCACCCAAAGTCTGGTGAAGGGCCCCCTACTTGTCTGCA 121

QY 83 GCAGGAGGCACATGCCAAGCTTGATAACTTGACCTTGATGGAGATCAA 130
Db 120 GGACGATGCCAGGCCAAGGTGGTCAACATCACCATCGTGCGCCCTCAA 73
```

RESULT 15  
US-09-864-761-789  
; Sequence 789, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aemlica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 789  
; LENGTH: 468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP000215.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 10  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.8  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 11  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.2  
US-09-864-761-789

Query Match 10.9%; Score 32.8; DB 9; Length 468;

Best Local Similarity 54.0%; Pred. No. 0.71;  
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 82 AGCAGGAGGCACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCGGGGA 141  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 106 AGCAGAGCGCCCAAGGACAACTTGCAGAGGAGTGGGTATCGAGCACTCACGCTGGGAGG 165  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 142 CTTTCTCTACACAAGCGCTCAACCCACATGTACAAACTCGGCACGACGAACTCCAGCCTCTCTGG 201  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 166 CTTTCTCTAGGAGCGGAGCCAGCTCTCTGAGCTCCAGAGGACCTTGGCGCTGAGA 225  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 202 AGAG 205  
Db 226 AGAG 229

Search completed: October 14, 2003, 17:17:33  
Job time : 148.333 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 .Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 11:32:16 ; Search time 1272.67 Seconds  
(without alignments)  
5729.185 Million cell updates/sec

Title: US-09-854-124-7  
Perfect score: 300  
Sequence: 1 atcatgcttcagacaacatc.....tggggattaaccttcagac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_estchum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_est1:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249.6	83.2	914	14	CA455226
2	245.4	81.8	709	10	BE561525
3	245.4	81.8	840	13	BU601226
4	244.8	81.6	624	10	BG284180
					CA455226 AGENCOURT
					BE561525 601345332
					BU601226 AGENCOURT
					BG284180 602408227

5	244.8	81.6	680	10	BE274253
6	244.8	81.6	736	10	BE312319
7	244.4	81.5	559	12	BM126492
8	244.4	81.5	601	12	BM043804
9	244.4	81.5	649	10	BE514071
10	244.4	81.5	659	13	BU729618
11	244.4	81.5	716	10	BE382866
12	244.4	81.5	801	12	BI196248
13	244.4	81.5	866	10	BF795157
14	244.4	81.5	876	10	BG181130
15	244.4	81.5	896	10	BE562088
16	244.4	81.5	914	10	BE547290
17	244.4	81.5	943	10	BF683514
18	244.4	81.5	1007	12	BM449472
19	244.4	81.5	1030	10	BF311745
20	244.4	81.5	1064	13	BQ277667
21	244.4	81.5	1071	13	BU184963
22	244.4	81.5	1182	11	AF125098
23	244.4	81.5	1201	9	AL582250
24	244.4	81.3	1201	9	AL582217
25	243.8	81.3	848	13	BU603101
26	243.8	81.1	833	13	BU597296
27	243	81.0	667	9	AU126087
28	242.8	80.9	488	14	CB160336
29	242.8	80.9	533	14	CB112523
30	242.8	80.9	696	9	AW249012
31	242.8	80.9	712	13	BU625683
32	242.8	80.9	714	10	BG104289
33	242.8	80.9	716	12	BG825252
34	242.8	80.9	730	10	BE296429
35	242.8	80.9	749	12	BG772776
36	242.8	80.9	790	10	BE795306
37	242.8	80.9	867	13	BQ233393
38	242.8	80.9	891	11	BC022839
39	242.8	80.9	921	10	BE799212
40	242.8	80.9	932	10	BE796384
41	242.8	80.9	1007	10	BE795838
42	242.8	80.9	1028	10	BE561044
43	241.8	80.6	909	13	BQ229290
44	241.2	80.4	628	10	BF035586
45	241.2	80.4	895	13	BU539659

#### ALIGNMENTS

RESULT 1

CA455226

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA455226 914 bp mRNA linear EST 12-NOV-2002  
AGENCOURT\_10714628 MAPcL Homo sapiens cDNA clone IMAGE:6722756 5',  
mRNA sequence.

CA455226

CA455226.1

GI:24905731

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 914)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Kristi A. Eglund, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM14285 row: p column: 20

High quality sequence stop: 553.

FEATURES  
source

Location/Qualifiers  
1. .914  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6722756"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1  
, LNCap"  
/lab\_host="EMDH10B"  
/clone\_lib="MAPcL"  
/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;  
Subtracted with brain, liver, lung, kidney and muscle.  
Directionally cloned. Priming method: oligo-dT. Average  
insert size: 1800 bp. Library amplification: 26,000 fold.  
Kristi A. Eglund, James J. Vincent, Robert Strausberg,  
Bungkok Lee & Ira Pastan: Discovery of new breast  
cancer genes encoding membrane and secreted proteins.  
Manuscript submitted."

BASE COUNT 246 a 233 c 246 g 188 t 1 others  
ORIGIN

Query Match 83.2%; Score 249.6; DB 14; Length 914;  
Best Local Similarity 94.1%; Pred. No. 1.9e-61;  
Matches 270; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 1 ATCATGCTTCAGACAACATCCGAGGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 60  
|||||  
Db 361 ATCATGCTTCAGACAACATCCGAGGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 420  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 120  
|||||  
Db 421 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 480  
QY 121 CACATGCCAAGCTGATTAACCTGACCTTGATGGAGATCAACACAGCGGGACTTTCCTCA 180  
|||||  
Db 481 CACATGCCAAGCTGATTAACCTGACCTTGATGGAGATCAACACAGCGGGACTTTCCTCA 540  
QY 181 CACAAGCGCTCAACACATGATCAAACTCCGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 240  
|||||  
Db 541 CACAAGCGCTCAACACATGATCAAACTCCGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 600  
QY 241 AG--CTAGGACTTCTAAGGAGGCTGTGTGAGCGGCTGTGTGGG 285  
|| |||||  
Db 601 AGTCTCAGGACTCTTANAAAAAGGCTGTGTGAGCGGCTGTGTGGG 647

RESULT 2  
BE561525  
LOCUS 601345332F1 NIH\_MGC\_8 Homo sapiens cDNA clone IMAGE:3678198 5',  
DEFINITION mRNA sequence.  
ACCESSION BE561525  
VERSION BE561525.1 GI:9805245  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLC354 row: p column: 07  
High quality sequence start: 16  
High quality sequence stop: 706.

FEATURES  
source

Location/Qualifiers  
1. .709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3678198"  
/tissue\_type="Burkitt lymphoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_8"  
/note="Organ: lymph; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5',  
adaptor: GGCACAG(S). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 185 a 178 c 203 g 142 t 1 others  
ORIGIN

Query Match 81.8%; Score 245.4; DB 10; Length 709;  
Best Local Similarity 94.8%; Pred. No. 2.8e-60;  
Matches 275; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACAACATCCGAGGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 60  
|||||  
Db 393 ATCATGCTTCAGACAACATCCGAGGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 452  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 120  
|||||  
Db 453 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 512  
QY 121 CACATGCCAAGCTGATTAACCTGACCTTGATGGAGATCAACACAGCGGGACTTTCCTCA 180  
|||||  
Db 513 CACATGCCAAGCTGATTAACCTGACCTTGATGGAGATCAACACAGCGGGACTTTCCTCA 572  
QY 181 CACAAGCGCTCAACACATGATCAAACTCCGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 240  
|||||  
Db 573 CACAAGCGCTCAACACATGATCAAACTCCGAGGAGGAAATCCGGACCTGGTCAAGGATATGT 632  
QY 241 AG--CTAGGACTTCTA-AAAAAGGCTGTGTGAGCGGCTGTGTGGGAT 287  
|| |||||  
Db 633 AGTCTCAGGACTTCTAGAGAAAGGCTGTGTGAGCGGCTGTGTGGGAT 682

RESULT 3  
BU601226 840 bp mRNA linear EST 20-SEP-2002  
LOCUS BU601226  
DEFINITION AGENCOURT\_10029782 NIH\_MGC\_142 Homo sapiens cDNA clone  
IMAGE:6495109 5', mRNA sequence.  
ACCESSION BU601226  
VERSION BU601226.1 GI:23252985  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: NCI  
CDNA Library Preparation: Michael Brownstein Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2672 row: g column: 14  
High quality sequence stop: 496.  
Location/Qualifiers

```

source
1..840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6495109"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="PH10B (T1-phase-resistant)"
/clone_lib="NIH_MGC_142"
/Note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggcgctcggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.3%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGCTGATCAACGAGAGTGGCCATTAGCGCGGG-3' and 5'-ATTCTAGAGCGGAGCGCGCCGACATG-DT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
BASE COUNT 217 a 215 c 221 g 185 t 2 others
ORIGIN

Query Match 81.8%; Score 245.4; DB 13; Length 840;
Best Local Similarity 93.0%; Pred. No. 3e-60;
Matches 267; Conservative 0; Mismatches 18; Indels 2; Gaps 1;

QY 1 ATCATGCTTCAGACACATCCGAGGCGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 60
DB 340 ATCATGCTTCAGACACATCCGAGGCGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 399
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTCGACAGCTTGTGAGACAGCAGGAGG 120
DB 400 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTCGACAGCTTGTGAGACAGCAGGAGG 459
QY 121 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCGGGGACTTTCTCTCA 180
DB 460 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCGGGGACTTTCTCTCA 519
QY 181 CACAAGCGCTCAACCATGTACAACTCCGACGAGAACCTCCAGCCTCTGGAAAGACCTC 240
DB 520 CACAAGCGCTCAACCATGTACAACTCCGACGAGAACCTCCAGCCTCTGGAGAGTACTC 579
QY 241 AG-CTAGGACTTCTAAAAAGCCCTGGTGTGAGCGCTTGGTTGGGG 285
DB 580 AGTCTCAGAGCTCTANAGAAGGCCCTGGTGTGAGCGGCTTGTGGGG 626

RESULT 4
BG284180 624 bp mRNA linear EST 21-FEB-2001
LOCUS 602408227F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520430 5',
DEFINITION mRNA sequence.
ACCESSION BG284180
VERSION BG284180.1 GI:13034868
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 624)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
found through the I.M.A.G.E. Consortium information can be

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10418 row: e column: 07
High quality sequence start: 4
High quality sequence stop: 621.
Location/Qualifiers
1..624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520430"
/tissue_type="adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/Note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 163 a 168 c 166 g 126 t 1 others
ORIGIN

Query Match 81.6%; Score 244.8; DB 10; Length 624;
Best Local Similarity 94.8%; Pred. No. 4e-60;
Matches 275; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCGAGGCGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 60
DB 333 ATCATGCTTCAGACACATCCGAGGCGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 392
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTCGACAGCTTGTGAGACAGCAGGAGG 120
DB 393 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTCGACAGCTTGTGAGACAGCAGGAGG 452
QY 121 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCGGGGACTTTCTCTCA 180
DB 453 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCGGGGACTTTCTCTCA 512
QY 181 CACAAGCGCTCAACCATGTACAACTCCGACGAGAACCTCCAGCCTCTGGAAAGACCTC 240
DB 513 CACAAGCGCTCAACCATGTACAACTCCGACGAGAACCTCCAGCCTCTGGAGAGTACTC 572
QY 241 AG-CTAGGACTTCTA-AAAAAGCCCTGGTGTGAGCGCTTGGTTGGGGAT 287
DB 573 AGTCTCAGAGCTTCTAGAGAAAGGCCCTGGTGTGAGCGGCTGTCTGGGGAT 622

RESULT 5
BE274253 680 bp mRNA linear EST 13-JUL-2000
LOCUS 601120527F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2967175 5',
DEFINITION mRNA sequence.
ACCESSION BE274253
VERSION BE274253.1 GI:9149193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 680)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

```

Plate: LLCM71 row: f column: 08  
High quality sequence stop: 651.

FEATURES source  
Location/Qualifiers  
1. .880  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2967175"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_20"  
/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 174 a 178 c 186 g 142 t  
ORIGIN

Query Match 81.6%; Score 244.8; DB 10; Length 680;  
Best Local Similarity 94.8%; Pred. No. 4.1e-60;  
Matches 275; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCGAGGCGAGCAGAAATCCGGACCTGTGTCAGGATATGT 60  
Db 318 ATCATGCTTCAGACACATCCGAGGCGAGCAGAAATCCGGACCTGTGTCAGGATATGT 377  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACAGCAGGAGG 120  
Db 378 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACAGCAGGAGG 437  
QY 121 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGGACTTTCCTCA 180  
Db 438 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGGACTTTCCTCA 497  
QY 181 CACAGCGCTCAACCATGATGACAACTCCGACGAACTCCAGCTCTGGAAGACCTC 240  
Db 498 CACAGCGCTCAACCATGATGACAACTCCGACGAACTCCAGCTCTGGAAGACCTC 557  
QY 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTGGGGAT 287  
Db 558 AGTCTCAGGACTTCTAGAGAAGCGCTGGTGCAGCGCTTGGTGGGGAT 607

RESULT 6  
BE312319  
LOCUS BE312319 736 bp mRNA linear EST 26-OCT-2000  
DEFINITION 601152780F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3508866 5', mRNA sequence.  
ACCESSION BE312319  
VERSION BE312319.1 GI:9130832  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 736)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCM190 row: p column: 19  
High quality sequence stop: 645.

## FEATURES source

Location/Qualifiers  
1. .736  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3508866"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_19"  
/note="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
Note: this is a NIH\_MGC library."

BASE COUNT  
ORIGIN

Query Match 81.6%; Score 244.8; DB 10; Length 736;  
Best Local Similarity 94.8%; Pred. No. 4.3e-60;  
Matches 275; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCGAGGCGAGCAGAAATCCGGACCTGTGTCAGGATATGT 60  
Db 367 ATCATGCTTCAGACACATCCGAGGCGAGCAGAAATCCGGACCTGTGTCAGGATATGT 426  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACAGCAGGAGG 120  
Db 427 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGAGACAGCAGGAGG 486  
QY 121 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGGACTTTCCTCA 180  
Db 487 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGGACTTTCCTCA 546  
QY 181 CACAGCGCTCAACCATGATGACAACTCCGACGAACTCCAGCTCTGGAAGACCTC 240  
Db 547 CACAGCGCTCAACCATGATGACAACTCCGACGAACTCCAGCTCTGGAAGACCTC 606  
QY 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTGGGGAT 287  
Db 607 AGTCTCAGGACTTCTAGAGAAGCGCTGGTGCAGCGCTTGGTGGGGAT 656

## RESULT 7

BM126492  
LOCUS BM126492 559 bp mRNA linear EST 12-MAR-2002  
DEFINITION if06h06.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens cDNA clone IMAGE:5675770 5', similar to TR:Q9Y248 Q9Y248 HSPC037 PROTEIN. ;, mRNA sequence.  
ACCESSION BM126492  
VERSION BM126492.1 GI:17121044  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 559)  
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y., and Bowers, Y.  
TITLE Endocrine Pancreas Consortium  
JOURNAL Unpublished  
COMMENT Other ESTs: if06h06.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812  
Fax: 617-495-8557

Email: dmelton@ohp.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brownefas.harvard.edu) This sequence now available from the IMAGE  
consortium, for clone orders contact: info@image.llnl.gov  
High quality sequence stop: 437.

#### FEATURES

Location/Qualifiers  
1..559  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5675770"  
/sex="Both"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/clone\_lib="Melton Normalized Human Islet 4 M4-HIS 1"  
/note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1;  
Site\_2: Sal 1; Starting library constructed using  
SuperScript plasmid library kit (Life Technologies). cDNA  
made by oligo-dT priming. Size-selected by column  
fractionation; average insert size 1.08 kb. Library was  
amplified once on solid support and plasmid DNA from  
library was prepared. The library DNA was normalized by  
method #4 from Bonaldo, Lennon, and Soares 1996 Genome  
Research 6:791-806; 0.5 microgram single-stranded library  
plasmid DNA was mixed with 5 micrograms PCR product  
representing library inserts and hybridized to an EcoT of  
20. Single-stranded (unhybridized) plasmids were isolated  
by hydroxyapatite chromatography and used to make this  
library."

BASE COUNT 157 a 131 c 142 g 129 t  
ORIGIN

Query Match 81.5%; Score 244.4; DB 12; Length 559;  
Best Local Similarity 95.1%; Pred. No. 4.9e-60;  
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCCGAGGACAGCAATCCGGACCTGGTCAAGGATATGT 60  
DB 88 ATCATGCTTCAGACACATCCCGAGGACAGCAATCCGGACCTGGTCAAGGATATGT 147  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120  
DB 148 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 207  
QY 121 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCAGCGGGACTTTCCCTCA 180  
DB 208 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCAGCGGGACTTTCCCTCA 267  
QY 181 CACAAGCGCTACACACATGTACAACTCCGACAGCACTCCAGCTCTGGAAAGACCTC 240  
DB 268 CACAAGCGCTACACACATGTACAACTCCGACAGCACTCCAGCTCTGGAGAGTACTC 327.  
QY 241 AG--CTPAGGACTTCTA-AAAAAGGCGCTGGTGCAGCGCTTGGTTGGG 285  
DB 328 AGTCTCAGGACTTCTAGAGAAAGGCGCTGGTGCAGGCGGCTTGTCTGGG 375

#### RESULT 8

BM043804  
LOCUS 603620658F1 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:5446215 5',  
DEFINITION mRNA sequence.  
ACCESSION BM043804  
VERSION BM043804.1 GI:16773071  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

#### REFERENCE

1 (bases 1 to 601)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1928 row: c column: 16  
High quality sequence stop: 598.

#### FEATURES

Location/Qualifiers  
1..601  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5446215"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTE7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 161 a 158 c 159 g 123 t  
ORIGIN

Query Match 81.5%; Score 244.4; DB 12; Length 601;  
Best Local Similarity 95.1%; Pred. No. 5.1e-60;  
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCCGAGGACAGCAATCCGGACCTGGTCAAGGATATGT 60  
DB 306 ATCATGCTTCAGACACATCCCGAGGACAGCAATCCGGACCTGGTCAAGGATATGT 365  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120  
DB 366 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 425  
QY 121 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCAGCGGGACTTTCCCTCA 180  
DB 426 CACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCAGCGGGACTTTCCCTCA 485  
QY 181 CACAAGCGCTACACACATGTACAACTCCGACAGCACTCCAGCTCTGGAAAGACCTC 240  
DB 486 CACAAGCGCTACACACATGTACAACTCCGACAGCACTCCAGCTCTGGAGAGTACTC 545  
QY 241 AG--CTPAGGACTTCTA-AAAAAGGCGCTGGTGCAGCGCTTGGTTGGG 285  
DB 546 AGTCTCAGGACTTCTAGAGAAAGGCGCTGGTGCAGGCGGCTTGTCTGGG 593

#### RESULT 9

BE514071  
LOCUS 601316576F1 NIH\_MGC\_8 Homo sapiens cDNA clone IMAGE:3635010 5',  
DEFINITION mRNA sequence.  
ACCESSION BE514071  
VERSION BE514071.1 GI:9721283  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 649)  
NH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Louis W. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LUCM330 row: h column: 19  
High quality sequence stop: 646.

Location/Qualifiers  
1. 649  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:3635010"  
/tissue\_type="Burkitt lymphoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_8"  
/notes="Organ: lymph; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dr priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GCCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)".

BASE COUNT 168 a 173 c 176 g 132 t  
ORIGIN

Query Match 81.5%; Score 244.4; DB 10; Length 649;  
Best Local Similarity 95.1%; Pred.No. 5.3e-60;  
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

```

QY   1 ATCATGCTTCAGACAACATCCGAAGCGACAGCAAAATCCGGACCCTGGTCAAGGATATGT 60
    |||||
Db   330 ATCATGCTTCAGACAACATCCGAAGCGACAGCAAAATCCGGACCCTGGTCAAGGATATGT 389
    |||||

QY   61 GGACACTCGTATAGCCAACTCGAGTGTCGTGACAGCTTTGTGAGACAGCAGGAGG 120
    |||||
Db   390 GGACACTCGTATAGCCAACTCGAGTGTCGTGACAGCTTTGTGAGACAGCAGGAGG 449
    |||||

QY   121 CACATGCCAAGCTGGGATAACTTGACCTTGATGGAGATCAACACAGCGGACATTCTCTCA 180
    |||||
Db   450 CACATGCCAAGCTGGGATAACTTGACCTTGATGGAGATCAACACAGCGGACATTCTCTCA 509
    |||||

QY   181 CACAAGCGCTCAACCCACATGTACAAAATCCGCGACGAACTCCAGCCCTCTGGAAGACCTC 240
    |||||
Db   510 CACAAGCGCTCAACCCACATGTACAAAATCCGCGACGAACTCCAGCCCTCTGGAAGACTC 569
    |||||

QY   241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGACGCGCTTGGTTGGGG 285
    || |||||||
Db   570 AGTCTCAGGACTTCTAGAGAAGCGCTGGTGACGCGCTTGGTTGGGG 617
    || |||||||

```

RESULT 10  
BU729618/c

LOCUS	BU729618
DEFINITION	UI-E-CQ1-ae-y-b-18-0-UI.s1 UI-E-CQ1 Homo sapiens cdna clone
SOURCE	UI-E-CQ1-ae-y-b-18-0-UI 3', mRNA sequence.
ACCESSION	BU729618
VERSION	BU729618.1 GI:23652681
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE 1 (bases 1 to 659)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Db 299 AGTCTCAGGACTTCTAGAGAAAGGCTGTGTGTCAGCGGCTTGTCTGGG 252

## RESULT 11

BE382866

LOCUS

DEFINITION

BE382866 716 bp mRNA linear EST 21-JUL-2000

60129775F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3627852 5',

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM311 row: n column: 13

High quality sequence stop: 651.

Location/Qualifiers

1. .716

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3627852"

/tissue\_type="neuroblastoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_19"

/notes="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2:

ECORI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT 185 a 179 c 195 g 157 t

ORIGIN

Query Match

Best Local Similarity 81.5%; Score 244.4; DB 10; Length 716;

Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCCGAAGGACAGCAAAATCCGGACCTTGTCAAGGATATGT 60

|||||

Db 307 ATCATGCTTCAGACACATCCCGAAGGACAGCAAAATCCGGACCTTGTCAAGGATATGT 366

|||||

QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120

|||||

Db 367 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 426

|||||

QY 121 CACATGCCAAGCTGGTAATCTTGACCTTGTATGGAGATCAACACCGAGCGGACTTTCTCTCA 180

|||||

Db 427 CACATGCCAAGCTGGTAATCTTGACCTTGTATGGAGATCAACACCGAGCGGACTTTCTCTCA 486

|||||

QY 181 CACAAGCGCTCAACACATGTACAACTCCGACGAACTCCAGCTTGTGGAAGACCTC 240

|||||

Db 487 CACAAGCGCTCAACACATGTACAACTCCGACGAACTCCAGCTTGTGGAAGACCTC 240

|||||

QY 241 AG-CTAGACTTCTA-AAAAAGGCTGTGTGTCAGCGCTTGTCTGGG 285

|||||

Db 547 AGTCTCAGGACTTCTAGAGAAAGGCTGTGTGTCAGCGGCTTGTCTGGG 594

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

## RESULT 12

BI196248

LOCUS

DEFINITION

BI196248 801 bp mRNA linear EST 10-JUL-2001

602794709F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4890147 5',

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM1767 row: b column: 04

High quality sequence stop: 781.

Location/Qualifiers

1. .801

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4890147"

/tissue\_type="neuroblastoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_19"

/notes="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2:

ECORI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT 208 a 196 c 216 g 181 t

ORIGIN

Query Match

Best Local Similarity 81.5%; Score 244.4; DB 12; Length 801;

Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCCGAAGGACAGCAAAATCCGGACCTTGTCAAGGATATGT 60

|||||

Db 326 ATCATGCTTCAGACACATCCCGAAGGACAGCAAAATCCGGACCTTGTCAAGGATATGT 385

|||||

QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120

|||||

Db 386 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 445

|||||

QY 121 CACATGCCAAGCTGGTAATCTTGACCTTGTATGGAGATCAACACCGAGCGGACTTTCTCTCA 180

|||||

Db 446 CACATGCCAAGCTGGTAATCTTGACCTTGTATGGAGATCAACACCGAGCGGACTTTCTCTCA 505

|||||

QY 181 CACAAGCGCTCAACACATGTACAACTCCGACGAACTCCAGCTTGTGGAAGACCTC 240

|||||

Db 506 CACAAGCGCTCAACACATGTACAACTCCGACGAACTCCAGCTTGTGGAAGACCTC 240

|||||

QY 241 AG-CTAGACTTCTA-AAAAAGGCTGTGTGTCAGCGCTTGTCTGGG 285

|||||

Db 566 AGTCTCAGGACTTCTAGAGAAAGGCTGTGTGTCAGCGGCTTGTCTGGG 613

```

RESULT 13
BF795157
LOCUS
DEFINITION 602256302F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4339537 5',
            mRNA sequence.
ACCESSION BF795157
VERSION   BF795157
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 866)
           NIH-MGC http://mgc.nci.nih.gov/
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9950 row: d column: 02
High quality sequence stop: 740.
FEATURES             location/Qualifiers
     source
     1..866
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:4339537"
     /tissue_type="lymphoma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_85"
     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
     Average insert size 1.867 Kb. Library enriched for
     full-length clones and constructed by Life Technologies.
     Note: this is a NIH_MGC Library."
BASE COUNT          220 a 221 c 236 g 189 t
ORIGIN
Query Match          81.5%; Score 244.4; DB 10; Length 866;
Best Local Similarity 95.1%; Pred. No. 6e-60;
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

Qy 1 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 60
    |||||
Db 367 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 426

Qy 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 120
    |||||
Db 427 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 486

Qy 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 180
    |||||
Db 487 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 546

Qy 181 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 240
    |||||
Db 547 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 606

Qy 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTTGGGG 285
    || |||||
Db 607 AGTCTCAGGACTTCTAGAGAAGCGCTTGGTGCAGCGCTTGGTTGGGG 654

RESULT 14
BG181130
LOCUS
DEFINITION 602329264F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4430823 5',
            mRNA sequence.
ACCESSION BG181130
VERSION   BG181130
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
           NIH-MGC http://mgc.nci.nih.gov/
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10184 row: o column: 16
High quality sequence stop: 674.
FEATURES             location/Qualifiers
     source
     1..876
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:4430823"
     /tissue_type="adenocarcinoma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_91"
     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
     Average insert size 1.4 Kb. Library enriched for
     full-length clones and constructed by Life Technologies.
     Note: this is a NIH_MGC Library."
BASE COUNT          246 a 217 c 214 g 199 t
ORIGIN
Query Match          81.5%; Score 244.4; DB 10; Length 876;
Best Local Similarity 95.1%; Pred. No. 6e-60;
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

Qy 1 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 60
    |||||
Db 85 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 144

Qy 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 120
    |||||
Db 145 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 204

Qy 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 180
    |||||
Db 205 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 264

Qy 181 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 240
    |||||
Db 265 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 324

Qy 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTTGGGG 285
    || |||||
Db 325 AGTCTCAGGACTTCTAGAGAAGCGCTTGGTGCAGCGCTTGGTTGGGG 372

RESULT 15
BE562088
LOCUS
DEFINITION 601345038F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678080 5',
            mRNA sequence.
ACCESSION BE562088
VERSION   BE562088
KEYWORDS  EST.

```

```

ACCESSION BG181130
VERSION   BG181130
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
           NIH-MGC http://mgc.nci.nih.gov/
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10184 row: o column: 16
High quality sequence stop: 674.
FEATURES             location/Qualifiers
     source
     1..876
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:4430823"
     /tissue_type="adenocarcinoma, cell line"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_91"
     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
     Average insert size 1.4 Kb. Library enriched for
     full-length clones and constructed by Life Technologies.
     Note: this is a NIH_MGC Library."
BASE COUNT          246 a 217 c 214 g 199 t
ORIGIN
Query Match          81.5%; Score 244.4; DB 10; Length 876;
Best Local Similarity 95.1%; Pred. No. 6e-60;
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

Qy 1 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 60
    |||||
Db 85 ATCATGCTTCAGACAACATCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGT 144

Qy 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 120
    |||||
Db 145 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGG 204

Qy 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 180
    |||||
Db 205 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 264

Qy 181 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 240
    |||||
Db 265 CACAAGCGCTCAACCAATGTACAACTCCGACGAACTCCAGCACTCTGGAAGACCTC 324

Qy 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTTGGGG 285
    || |||||
Db 325 AGTCTCAGGACTTCTAGAGAAGCGCTTGGTGCAGCGCTTGGTTGGGG 372

RESULT 15
BE562088
LOCUS
DEFINITION 601345038F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678080 5',
            mRNA sequence.
ACCESSION BE562088
VERSION   BE562088
KEYWORDS  EST.

```



SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 896)  
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps@email.nih.gov](mailto:cgaps@email.nih.gov)  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LCM354 row: k column: 09  
 High quality sequence start: 24  
 High quality sequence stop: 840.  
 Location/Qualifiers  
 1..896  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3678080"  
 /tissue\_type="Burkitt lymphoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC 8"  
 /notes="Organ: lymph; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 218 a 230 c 251 g 197 t  
 ORIGIN  
 Query Match 81.5%; Score 244.4; DB 10; Length 896;  
 Best Local Similarity 95.1%; Pred. No. 6.1e-60;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;  
 QY 1 ATCATGCTTCAGACACATCCCGAAGGCAGACGAAATCCGGACCCCTGGTCAAGGATATGT 60  
 Db |||||  
 QY 408 ATCATGCTTCAGACACATCCCGAAGGCAGACGAAATCCGGACCCCTGGTCAAGGATATGT 457  
 Db |||||  
 QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACACGAGG 120  
 Db |||||  
 QY 468 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACACACGAGG 527  
 Db |||||  
 QY 121 CACATGCCAAGCTGGATTAACCTTGATGGAGATCAACACCGGGGACTTTCCTCA 180  
 Db |||||  
 QY 528 CACATGCCAAGCTGGATTAACCTTGATGGAGATCAACACCGGGGACTTTCCTCA 587  
 Db |||||  
 QY 181 CACAAGCGCTCAACCATGTACAACTCCGACGAACTCCAGCCCTCTGGAAGACCTC 240  
 Db |||||  
 QY 588 CACAAGCGCTCAACCATGTACAACTCCGACGAACTCCAGCCCTCTGGAAGTACTC 647  
 QY 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTGGG 285  
 Db || |||||  
 QY 648 AGTCTCAGGACTTCTAGAGAAAGGCGCTGGTGCAGCGCTTGGTGGG 695

Search completed: October 14, 2003, 14:09:10  
 Job time : 1273.67 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 14:17:11 ; Search time 143.333 Seconds  
(without alignments)  
5431.000 Million cell updates/sec

Title: US-09-854-124-7

Perfect score: 300

Sequence: 1 atcatgttcagacaacatc.....tggggattaaccccttcagac 300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/ECT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/ECTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	300	100.0	300	9	US-09-854-124-7
2	244.4	81.5	1210	9	Sequence 7, Appl1
3	237.2	79.1	484	11	Sequence 214, App
4	224	74.7	597	10	Sequence 19338, A
5	224	74.7	597	10	Sequence 4407, Ap
6	214.4	71.5	300	9	Sequence 4407, Ap
7	199.2	66.4	382	14	Sequence 5, Appl1
8	199.2	66.4	382	14	Sequence 9460, Ap
9	39.2	13.1	820	13	Sequence 9460, Ap
10	39.2	13.1	820	13	Sequence 151845, Ap
11	35	11.7	481	11	Sequence 151846, Ap
12	33.4	11.1	476	11	Sequence 870, App
13	32.6	10.9	39061	12	Sequence 2536, Ap
14	32.6	10.9	39061	12	Sequence 148, App
15	31.8	10.6	468	9	Sequence 254, App
16	31.8	10.6	737	9	Sequence 789, App
					Sequence 17572, A

17	31.8	10.6	814	13	US-10-027-632-156310
18	31.6	10.5	650	13	Sequence 156310, Sequence 246523, Sequence 1387, Ap
19	31.4	10.5	517	10	US-09-998-598-1387
20	31.4	10.5	596	13	Sequence 50355, A
21	31.4	10.5	596	13	Sequence 50356, A
22	31.4	10.5	596	13	Sequence 50357, A
23	31.2	10.4	546	10	Sequence 2478, Ap
24	31.2	10.4	3309400	10	Sequence 1, Appl1
25	31	10.3	408	10	US-09-960-352-14278
26	31	10.3	434	10	Sequence 12424, A
27	31	10.3	717	14	Sequence 43, Appl1
28	31	10.3	1623	14	Sequence 33, Appl1
29	31	10.3	1827	11	Sequence 608, App
30	31	10.3	1827	14	Sequence 608, App
31	31	10.3	2439	14	Sequence 3, Appl1
32	31	10.3	2439	14	Sequence 5, Appl1
33	31	10.3	3171	14	Sequence 31, Appl1
34	31	10.3	4833	14	Sequence 21, Appl1
35	31	10.3	4944	13	Sequence 51, Appl1
36	31	10.3	4951	13	Sequence 56, Appl1
37	31	10.3	6811	11	Sequence 11, Appl1
38	31	10.3	8999	13	Sequence 48, Appl1
39	30.8	10.3	487	9	Sequence 291, App
40	30.8	10.3	558	9	Sequence 7880, App
41	30.8	10.3	811	13	Sequence 140232, Sequence 5598, App
42	30.8	10.3	1371	14	Sequence 111, App
43	30.8	10.3	3051	13	Sequence 1, Appl1
44	30.8	10.3	9025608	14	Sequence 19581, A
45	30.4	10.1	476	11	Sequence 19581, A

#### ALIGNMENTS

#### RESULT 1

US-09-854-124-7  
; Sequence 7, Application US/09854124  
; Patent No. US20020076735A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Innis, Michael A.  
; APPLICANT: Garcia, Pablo Dominiguez  
; APPLICANT: Sudduth-Klinger, Julie  
; APPLICANT: Reinhard, Christoph  
; APPLICANT: Giese, Klaus  
; APPLICANT: Randazzo, Filippo  
; APPLICANT: Kennedy, Giulia C.  
; APPLICANT: Pot, David  
; APPLICANT: Kassam, Altaf  
; APPLICANT: Lamson, George  
; APPLICANT: Drmanac, Radoje  
; APPLICANT: Crkvenjakov, Radomir  
; APPLICANT: Dickson, Mark  
; APPLICANT: Drmanac, Shezana  
; APPLICANT: Labat, Ivan  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Kita, David  
; APPLICANT: Jones, Lee William  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Stache-Crain, Birgit  
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods Using Molecules Differentially Expressed in Cancer Cells  
; FILE REFERENCE: 2300-1490  
; CURRENT APPLICATION NUMBER: US/09/854,124  
; CURRENT FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: 09/400,947  
; PRIOR FILING DATE: 1999-09-22  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 300  
; TYPE: DNA

RESULT 4  
US-09-796-692-4407  
Sequence 4407, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01

;; PRIOR APPLICATION NUMBER: 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 4407  
;; LENGTH: 597  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (583)  
;; NAME/KEY: unsure  
;; LOCATION: (592)  
;; OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-4407

Query Match 74.7%; Score 224; DB 10; Length 597;  
Best Local Similarity 99.1%; Pred. No. 4.7e-67;  
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATCATGCTTCAGACAAACATCCGAGGCGAGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 60  
Db 372 ATCATGCTTCAGACAAACATCCGAGGCGAGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 431  
QY 61 GGGACACTCGTATAGCCAAACCTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120  
Db 432 GGGACACTCGTATAGCCAAACCTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 491  
QY 121 CATATGCCAAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGAGCTTTCCCTCA 180  
Db 492 CATATGCCAAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGAGCTTTCCCTCA 551  
QY 181 CACAGCGCTCAACACATGTACAAACTCCGACGAACTCCAGCC 226  
Db 552 CACAAGCGCTCAACACATGTACAAACTCCGACGAACTCCAGCC 597

RESULT 5  
US-10-040-862-4407  
;; Sequence 4407, Application US/10040862  
;; Publication No. US20030078396A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Retter, Marc  
;; APPLICANT: Corixa Corporation  
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
;; FILE REFERENCE: 014058-013520US  
;; CURRENT APPLICATION NUMBER: US/10/040,862  
;; CURRENT FILING DATE: 2001-11-06  
;; PRIOR APPLICATION NUMBER: US 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: US 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: US 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 60/200,779  
;; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 4407  
;; LENGTH: 597  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (583)  
;; OTHER INFORMATION: n=A,T,C or G  
;; NAME/KEY: unsure  
;; LOCATION: (592)  
;; OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-4407

Query Match 74.7%; Score 224; DB 14; Length 597;  
Best Local Similarity 99.1%; Pred. No. 4.7e-67;  
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATCATGCTTCAGACAAACATCCGAGGCGAGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 60  
Db 372 ATCATGCTTCAGACAAACATCCGAGGCGAGAGAGAAATCCGGACCCCTGGTCAAGGATATGT 431  
QY 61 GGGACACTCGTATAGCCAAACCTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120  
Db 432 GGGACACTCGTATAGCCAAACCTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 491  
QY 121 CATATGCCAAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGAGCTTTCCCTCA 180  
Db 492 CATATGCCAAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCGCGGAGCTTTCCCTCA 551  
QY 181 CACAGCGCTCAACACATGTACAAACTCCGACGAACTCCAGCC 226  
Db 552 CACAAGCGCTCAACACATGTACAAACTCCGACGAACTCCAGCC 597

RESULT 6  
US-09-854-124-5  
;; Sequence 5, Application US/09854124  
;; Patent No. US20020076735A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Williams, Lewis T.  
;; APPLICANT: Escobedo, Jaime  
;; APPLICANT: Innis, Michael A.  
;; APPLICANT: Garcia, Pablo Dominiguez  
;; APPLICANT: Sudduth-Klinger, Julie  
;; APPLICANT: Reinhard, Christoph  
;; APPLICANT: Giese, Klaus  
;; APPLICANT: Randazzo, Filippo  
;; APPLICANT: Kennedy, Giulia C.  
;; APPLICANT: Pot, David  
;; APPLICANT: Kassam, Altat  
;; APPLICANT: Lamson, George  
;; APPLICANT: Drmanac, Radoje  
;; APPLICANT: Crkvenjakov, Radomir  
;; APPLICANT: Dickson, Mark

```
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods Using
; TITLE OF INVENTION: Molecules Differentially Expressed in Cancer Cells
; FILE REFERENCE: 2300-1490
; CURRENT APPLICATION NUMBER: US/09/854,124
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/400,947
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
;
; Query Match          71.5%; Score 214.4; DB 9; Length 300;
; Best Local Similarity 94.6%; Pred. No. 7.5e-64;
; Matches 244; Conservative 0; Mismatches 11; Indels 3; Gaps 2;
;
QY 31 ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACATCTGATAGCCAAACTCCGAGTGT 90
Db 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACATCTGATAGCCAAACTCCGAGTGT 60
QY 91 CTGCTGACAGCTTTGTGAGACAGAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 150
Db 61 CTGCTGACAGCTTTGTGAGACAGAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 120
QY 151 TGGAGATCAACACAGCGGGGACTTTCCTCACACAGGCTCAACCATGTACAACTCC 210
Db 121 TGGAGATCAACACAGCGGGGACTTTCCTCACACAGGCTCAACCATGTACAACTCC 180
QY 211 GCAGCAACTCCAGCCCTCTGGAAGACCTCAG--CTAGGACTTCTA-AAAAAGCCCTGGT 267
Db 181 GCAGCAACTCCAGCCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240
QY 268 GCAGCCGCTTGGTTGGGG 285
Db 241 GCAGCCGCTTGGTTGGGG 258

RESULT 7
US-09-796-692-9460/c
; Sequence 9460, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
;
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9460
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (72)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (80)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (89)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (95)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (122)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (129)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (131)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (155)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (173)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (222)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (235)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (236)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (237)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (245)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (251)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (305)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
```









```
RESULT 14
US-10-085-959-254/c
; Sequence 254, Application US/10085959
; Publication No. US20030165870A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Welch, Rodney A.
; APPLICANT: Burland, Valerie D.
; TITLE OF INVENTION: No. US20030165870A1el Sequence of E. Coli CFT073
; FILE REFERENCE: 960296.97648
; CURRENT APPLICATION NUMBER: US/10/085,959
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/242,412
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 254
; LENGTH: 39061
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (35920)..(35920)
; OTHER INFORMATION: Unsure
; NAME/KEY: CDS
; LOCATION: (31941)..(34667)
; OTHER INFORMATION:
; US-10-085-959-254

Query Match      10.9%; Score 32.6; DB 12; Length 39061;
Best Local Similarity 51.0%; Pred. No. 3;
Matches 77; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 88 TGCTGTGTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATACTTGACCT 147
Db 15300 TGGAGATTACACCTTTGAAGTTTACCACGACGCGGAGCGGTGGGACCAAAACC 15241

Qy 148 TGATGGAGATCAACACAGCGGGGACTTCCTCACACAGAGCGCTCAACCATGTGACAAAC 207
Db 15240 TGTTGGGTATTACCACAGCGGTGCATGCGATGGCAGAGCTGTGTGAACCTACGTGTTAAAA 15181

Qy 208 TCCGACGAACCTCCAGCCTCTGGAAGACC 238
Db 15180 TTATCGAAGCAACCTGCTGCCGAATACCC 15150

RESULT 15
US-09-864-761-789
; Sequence 789, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 789
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000215.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.2
US-09-864-761-789

Query Match      10.6%; Score 31.8; DB 9; Length 468;
Best Local Similarity 51.8%; Pred. No. 1.1;
Matches 72; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 112 AGCAGGAGGCACATGCCAAGCTGGATTAACCTTGATGAGATCAACACAGCGGGA 171
Db 106 AGCAGCGGCCAAGGACAACTTCGAGAAGGAGCTGCGTATCGAGCACTACGCTGCGAGG 165

Qy 172 CTTTCTCTCACACAAGCGCTCAACCCACATGTACAAACTCCGACCAACCTCCAGGCTCTGG 231
Db 166 CTTGCTGGCTCAGAGCGGAGCGAGCTCTCTGAGCTCCAGAGGAGCTTCGCGGCTGAGA 225

Qy 232 AAAGACCTCAGCTAGGACT 250
Db 226 AGAGCCGCACTCGGAGCT 244

Search completed: October 14, 2003, 17:17:50
Job time : 149.333 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:05:16 ; Search time 36.6667 Seconds  
(without alignments)  
3611.317 Million cell updates/sec

Title: US-09-854-124-7  
Perfect score: 300  
Sequence: 1 atcatgcttcagacaacatc.....tggggattaaccttcagac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	11.7	2241	2	US-08-838-219B-20
2	35	11.7	2241	3	US-09-233-336A-20
3	35	11.7	2241	3	US-09-233-752A-20
4	35	11.7	2241	3	US-09-402-036-20
5	35	11.7	2241	4	US-09-904-226-20
6	35	11.7	2370	2	US-08-838-219B-19
7	35	11.7	2370	3	US-09-233-336A-19
8	35	11.7	2370	3	US-09-233-752A-19
9	35	11.7	2370	3	US-09-402-036-19
10	35	11.7	2370	4	US-09-904-226-19
11	35	11.7	2403	1	US-08-471-033-30
12	35	11.7	2403	2	US-08-471-044-30
13	35	11.7	2403	2	US-08-463-483A-30
14	35	11.7	2403	2	US-08-471-046A-30
15	35	11.7	2403	2	US-08-470-568B-30
16	35	11.7	2403	2	US-08-838-219B-7
17	35	11.7	2403	2	US-08-469-334-30
18	35	11.7	2403	3	US-09-300-529-30
19	35	11.7	2403	3	US-09-233-336A-7
20	35	11.7	2403	3	US-09-233-752A-7
21	35	11.7	2403	3	US-09-402-036-7
22	35	11.7	2403	4	US-09-904-226-7
23	32.6	10.9	1296	4	US-09-252-991A-7149
24	32.6	10.9	1419	4	US-09-252-991A-7347
25	32.6	10.9	2748	4	US-09-252-991A-7208
26	31.6	10.5	408	4	US-09-252-991A-9970
c 27	31.6	10.5	900	4	US-09-252-991A-10183

28 31.6 10.5 1332 4 US-09-252-991A-9889 Sequence 9889, Ap  
29 31 10.3 717 4 US-09-513-783A-43 Sequence 43, Appl  
30 31 10.3 1623 4 US-09-513-783A-33 Sequence 33, Appl  
31 31 10.3 2439 4 US-09-513-783A-3 Sequence 3, Appl  
32 31 10.3 2439 4 US-09-513-783A-5 Sequence 5, Appl  
33 31 10.3 3171 4 US-09-513-783A-31 Sequence 31, Appl  
34 31 10.3 4833 4 US-09-513-783A-21 Sequence 21, Appl  
35 30.8 10.3 1625 2 US-08-709-923-2 Patent No. 5472691  
36 29.8 9.9 666 6 5472691-4 Patent No. 5472691  
37 29.8 9.9 669 3 US-08-556-965-1 Patent No. 5472691  
38 29.8 9.9 723 6 5472691-7 Patent No. 5472691  
39 29.8 9.9 1389 2 US-08-023-980B-3 Sequence 3, Appl  
40 29.8 9.9 1389 2 US-08-486-953A-3 Patent No. 5472691  
41 29.8 9.9 1396 6 5472691-1 Patent No. 5472691  
42 29.8 9.9 10079 2 US-08-476-866-20 Sequence 20, Appl  
c 43 29.6 9.9 4403765 3 US-09-103-840A-2 Sequence 2, Appl  
c 44 29.6 9.9 4411529 3 US-09-103-840A-1 Sequence 1, Appl  
45 29.4 9.8 717 4 US-09-513-783A-49 Sequence 49, Appl

## ALIGNMENTS

RESULT 1  
US-08-838-219B-20  
; Sequence 20, Application US/08838219B  
; Patent No. 5877012  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Gregory W  
; APPLICANT: Kozziel, Michael G  
; APPLICANT: Mullins, Martha A  
; APPLICANT: Nye, Gordon J  
; APPLICANT: Carr, Brian  
; APPLICANT: Desai, Nalini M  
; APPLICANT: Kostichka, N. Kristy  
; APPLICANT: Duck, Nicholas B  
; APPLICANT: Estruch, Juan J  
; TITLE OF INVENTION: A No. 5877012el Class of Proteins for the  
; TITLE OF INVENTION: Control of Plant Pests  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/838,219B  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/463,483  
; FILING DATE: 06-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/314,594  
; FILING DATE: 09-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/218,018  
; FILING DATE: 23-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/037,057  
; FILING DATE: 25-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Gary M.  
; REGISTRATION NUMBER: 40,403  
; REFERENCE/DOCKET NUMBER: CGC 1925  
; TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA encoding
; DESCRIPTION: VIP3A(c)"
; HYPOTHETICAL: NO
;
US-08-838-219B-20

Query Match 11.7%; Score 35; DB 2; Length 2241;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACAGCAGGAGCAGATGCCAAGCTGGGATAACTTGACCTTGATGGAGATCAACACGAG 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1950 GAGCAGACAGCGGACGAGCGCTGGGGCGACAACCTTCATCATCTCGGAGATCAGCCGAG 2009

QY 167 CGGACTTTCTCTCACACAAGCGCTCAACACCATGTACAA 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2010 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2048

RESULT 2
US-09-233-336A-20
; Sequence 20, Application US/09233336A
; Patent No. 6107279
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: A No. 6107279el Class of Proteins for the
; TITLE OF INVENTION: Control of Plant Pests
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,336A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,219
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
```

```
;
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1925
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "synthetic DNA encoding
; DESCRIPTION: VIP3A(c)"
; HYPOTHETICAL: NO
;
US-09-233-336A-20

Query Match 11.7%; Score 35; DB 3; Length 2241;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACAGCAGGAGCAGATGCCAAGCTGGGATAACTTGACCTTGATGGAGATCAACACGAG 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1950 GAGCAGACAGCGGACGAGCGCTGGGGCGACAACCTTCATCATCTCGGAGATCAGCCGAG 2009

QY 167 CGGACTTTCTCTCACACAAGCGCTCAACACCATGTACAA 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2010 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2048

RESULT 3
US-09-233-752A-20
; Sequence 20, Application US/09233752A
; Patent No. 6137033
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: A No. 6137033el Class of Proteins for the
; TITLE OF INVENTION: Control of Plant Pests
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/233,752A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,219
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
```

```

; GENERAL INFORMATION:
; APPLICANT: Estruch, Juan J.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Desai, Nalini
; APPLICANT: Kozziel, Michael
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: Plant Pest Control
; FILE REFERENCE: S-21284D
; CURRENT APPLICATION NUMBER: US/09/904,226
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/EP98/01952
; PRIOR FILING DATE: 1998-04-02

```

Query Match	11.7%	Score 35;	DB 4;	Length 2241;
Best Local Similarity	59.6%	Pred. No. 0.14;		
Matches	59;	Conservative 0;	Mismatches 40;	Indels 0; Gaps 0;
QY	107	GAGACAGCAGGAGGCACATGCCAAGCTGGATACTTGACCTTGATGGAGATCAACACCAG	166	
Db	1950	GAGCCAGAGCGGCGAGCGGCCCTGGGCGCAACAATTCATCATCCTGGAGATCAGCCGGAG	2009	
QY	167	CGGGAGCTTTCCTCACACAAGCGCTCAACCACATGTACAA	205	
Db	2010	CGAAGAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA	2048	

; Sequence 19, Application US/08838219B  
; Patent No. 5877012

## ; GENERAL INFORMATION:

; APPLICANT: Warren, Gregory W  
; APPLICANT: Koziel, Michael G  
; APPLICANT: Mullins, Martha A  
; APPLICANT: Nye, Gordon J  
; APPLICANT: Carr, Brian  
; APPLICANT: Desai, Nalini M  
; APPLICANT: Kostichka, N. Kristy  
; APPLICANT: Duck, Nicholas B  
; APPLICANT: Estruch, Juan J

; TITLE OF INVENTION: A No. 5877012el Class of Proteins for the  
; TITLE OF INVENTION: Control of Plant Pests

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CIBA-GEIGY Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/838,219B

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/463,483

; FILING DATE: 06-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/314,594

; FILING DATE: 09-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/218,018

; FILING DATE: 23-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/037,057

; FILING DATE: 25-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Pace, Gary M.

; REGISTRATION NUMBER: 40,403

; REFERENCE/DOCKET NUMBER: CGC 1925

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8582

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2370 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "synthetic DNA encoding

; DESCRIPTION: VIP3A(b)"

; HYPOTHETICAL: NO

; US-08-838-219B-19

Query Match 11.7%; Score 35; DB 2; Length 2370;

Best Local Similarity 59.6%; Pred. No. 0.14;

Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACACGAGGCGACATCGCAGCTGGGATTAACCTTGATGATGATCAACACCAG 166

Db 1950 GAGCCAGACGGCAGCGGCGCTGGGCGGACACTTCATCTCTGGAGATCAACCCGAG 2009

QY 167 CGGAGACTTCTCTACACAAGCGCTCAACACCATGTACAA 205

Db 2010 CGAGAACTGCTGAGCCCGGAGCTGTATCAACACCCACAA 2048

## RESULT 7

US-09-233-336A-19

; Sequence 19, Application US/09233336A

; Patent No. 6107279

; GENERAL INFORMATION:

; APPLICANT: Warren, Gregory W

; APPLICANT: Koziel, Michael G

; APPLICANT: Mullins, Martha A

; APPLICANT: Nye, Gordon J

; APPLICANT: Carr, Brian

; APPLICANT: Desai, Nalini M

; APPLICANT: Kostichka, N. Kristy

; APPLICANT: Duck, Nicholas B

; APPLICANT: Estruch, Juan J

; TITLE OF INVENTION: A No. 6107279el Class of Proteins for the

; TITLE OF INVENTION: Control of Plant Pests

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CIBA-GEIGY Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/233,336A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/838,219

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/314,594

; FILING DATE: 09-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/218,018

; FILING DATE: 23-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/037,057

; FILING DATE: 25-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Pace, Gary M.

; REGISTRATION NUMBER: 40,403

; REFERENCE/DOCKET NUMBER: CGC 1925

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8582

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2370 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "synthetic DNA encoding

; DESCRIPTION: VIP3A(b)"

; HYPOTHETICAL: NO

; US-09-233-336A-19

Query Match 11.7%; Score 35; DB 3; Length 2370;

Best Local Similarity 59.6%; Pred. No. 0.14;

Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACACGAGGCGACATCGCAGCTGGGATTAACCTTGATGATGATCAACACCAG 166

Db 1950 GAGCCAGACGGCAGCGGCGCTGGGCGGACAACTTCATCTCTGGAGATCAACCCGAG 2009



```
; APPLICANT: Koziel, Michael
; APPLICANT: Nye, Gordon
; TITLE OF INVENTION: Plant Pest Control
; FILE REFERENCE: S-21284D
; CURRENT APPLICATION NUMBER: US/09/904,226
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/EP98/01952
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 08/838,219
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,263
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/832,265
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: 08/463,483
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/314,594
; PRIOR FILING DATE: 1994-09-09
; PRIOR APPLICATION NUMBER: 08/218,018
; PRIOR FILING DATE: 1994-03-24
; PRIOR APPLICATION NUMBER: 08/037,057
; PRIOR FILING DATE: 1993-03-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
; OTHER INFORMATION: encoding VIP3A(b)
; US-09-904-226-19

Query Match          11.7%; Score 35; DB 4; Length 2370;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 107 GAGACAGCAGGCGACATGCCAAGCTGGGATGAACCTTGACCTTGATGGAGATCAACACCG 166
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1950 GAGCGACAGGCGGACGACGAGGCTGGGCGGACAACTTCATCATCTCGGAGATCAGCCGCG 2009

Qy 167 CGGACATTTCTCACACAAGCGCTCAACACCATGTACAA 205
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2010 CGAGAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2048

RESULT 11
; Sequence 30, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SQLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA"
; OTHER INFORMATION: sequence encoding VIP3A(a)"
; US-08-471-033-30

Query Match          11.7%; Score 35; DB 1; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 107 GAGACAGCAGGCGACATGCCAAGCTGGGATGAACCTTGACCTTGATGGAGATCAACACCG 166
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1966 GAGCGACAGGCGGACGAGGCGCTGGGCGGACAACTTCATCATCTCGGAGATCAGCCGCG 2025

Qy 167 CGGACATTTCTCACACAAGCGCTCAACACCATGTACAA 205
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2026 CGAGAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064

RESULT 12
US-08-471-044-30
; Sequence 30, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
```



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,044  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/463,483  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594  
FILING DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018  
FILING DATE: 23-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/037,057  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40,403  
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLv3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8582  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2403 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Synthetic DNA"  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 11..2389  
OTHER INFORMATION: /note= "maize optimized DNA"  
OTHER INFORMATION: sequence encoding VIP3A(a)"  
US-08-471-044-30

Query Match 11.7%; Score 35; DB 2; Length 2403;  
Best Local Similarity 59.6%; Pred. No. 0.14;  
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 107 GAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCG 166  
DB 1966 GAGCCAGAACGCGGAGGCGCTGGGGCGGACAACTTCATCTCTGGAGATCAGCCCGAG 2025  
QY 167 CGGGACTTTCCTCACACAAAGCGCTCAACACATGTACAA 205  
DB 2026 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064

RESULT 13  
US-08-463-483A-30  
Sequence 30, Application US/08463483A  
Patent No. 5849870  
GENERAL INFORMATION:  
APPLICANT: Warren, Gregory W  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalini M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains  
NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,483A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594  
FILING DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018  
FILING DATE: 23-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/037,057  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2403 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Synthetic DNA"  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 11..2389  
OTHER INFORMATION: /note= "maize optimized DNA"  
OTHER INFORMATION: sequence encoding VIP3A(a)"  
US-08-463-483A-30

Query Match 11.7%; Score 35; DB 2; Length 2403;  
Best Local Similarity 59.6%; Pred. No. 0.14;  
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
QY 107 GAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACCG 166  
DB 1966 GAGCCAGAACGCGGAGGCGCTGGGGCGGACAACTTCATCTCTGGAGATCAGCCCGAG 2025  
QY 167 CGGGACTTTCCTCACACAAAGCGCTCAACACATGTACAA 205  
DB 2026 CGAGAAGCTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064

RESULT 14  
US-08-471-046A-30  
Sequence 30, Application US/08471046A  
Patent No. 5866326  
GENERAL INFORMATION:  
APPLICANT: Warren, Gregory W  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalini M  
APPLICANT: Kostichka, N. Kristy

```
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal
; TITLE OF INVENTION: Protein Genes
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 586326artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,046A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SOLv4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA"
; OTHER INFORMATION: sequence encoding VIP3A(a)"
;
; US-08-471-046A-30
;
Query Match 11.7%; Score 35; DB 2; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACACGAGGAGGACATGCCAAGCTGGATACTTGACCTTGATGATGAGATCAACACG 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1966 GAGCCAGACGGCGACGAGGCTGGGGCGACACTTCATCATCTCGGATCAGCCGCG 2025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 167 CGGAGCTTTCTCACAACGAGCTCAACACCATGTACAA 205
    || || || || || || || || || || || || || || || || || || || || ||
Db 2026 CGAAGACTGCTGAGCCCGGAGCTGATCAACACCAACAA 2064
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 15
; Sequence 30, Application US/08470566B
; Patent No. 5872212
```

```
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5872212el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5872212artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,566B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV4 - SOLv4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2403 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11..2389
; OTHER INFORMATION: /note= "maize optimized DNA"
; OTHER INFORMATION: sequence encoding VIP3A(a)"
;
; US-08-470-566B-30
;
Query Match 11.7%; Score 35; DB 2; Length 2403;
Best Local Similarity 59.6%; Pred. No. 0.14;
Matches 59; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 107 GAGACACGAGGAGGACATGCCAAGCTGGATACTTGACCTTGATGATGAGATCAACACG 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1966 GAGCCAGACGGCGACGAGGCTGGGGCGACACTTCATCATCTCGGATCAGCCGCG 2025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 167 CGGAGCTTTCTCACAACGAGCTCAACACCATGTACAA 205
    || || || || || || || || || || || || || || || || || || || || ||
```

Db 2026 CGAGAGCTGCTGAGCCCGAGCTGATCAACACCAACAA 2064

Search completed: October 14, 2003, 17:04:30  
Job time : 38.6667 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 12:04:45 ; Search time 157.333 Seconds  
(without alignments)  
5147.234 Million cell updates/sec

Title: US-09-854-124-7  
Perfect score: 300  
Sequence: 1 atcatgttcagacaacatc.....tggggattaacaccttcagac 300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	24	Differentially exp
2	244.4	81.5	858	22	Human cDNA sequenc
3	244.4	81.5	1171	25	Human NF-kappaB as
4	244.4	81.5	1210	21	Human cancer assoc
5	243	81.0	667	22	Human cDNA clone (
6	214.4	71.5	300	24	Differentially exp
7	205.4	68.5	678	23	DNA encoding novel
8	205.4	68.5	678	23	DNA encoding novel

9	194	64.7	1190	23	AAS87106	DNA encoding novel
10	194	64.7	1499	23	AAS87104	DNA encoding novel
11	158	52.7	454	24	ABL68902	Kidney cancer rela
12	130.6	43.5	603	24	ABO59992	Human colon cancer
13	75.2	25.1	649	24	ABO59972	Human colon cancer
14	59.6	19.9	612	23	ABL13607	Drosophila melanog
15	59.6	19.9	2870	23	ABL19606	Drosophila melanog
16	59.6	19.9	4904	23	ABL05636	Drosophila melanog
17	44.6	14.9	51	22	AAL27661	Human SNP oligonuc
18	42	14.0	1257	23	AAS83051	DNA encoding novel
19	40	13.3	940	23	AAS72563	DNA encoding novel
20	40	13.3	1143	25	ABX71197	Novel human cDNA s
21	39.6	13.2	734	22	AAH08244	Human cDNA clone (
22	39.6	13.2	3071	22	AAH18628	Human cDNA sequenc
23	39	13.0	50	22	AAL34480	Human SNP oligonuc
24	36	12.0	654	24	ABN61061	Human cancer relat
25	35.8	11.9	1935	22	AAF60959	P. putida KT2440-a
26	35	11.7	2241	19	AAV68070	Maize-optimised DN
27	35	11.7	2370	19	AAV68069	Maize-optimised DN
28	35	11.7	2403	17	AAT13955	Maize-optimised VI
29	35	11.7	2403	18	AAV74004	Maize-optimised-B.
30	35	11.7	2403	19	AAV16177	Maize-optimised DN
31	35	11.7	2403	19	AAV16177	Maize-optimised DN
32	35	11.7	2444	17	AAT13964	VIP3A(a) synthetic
33	35	11.7	2444	18	AAT74011	B. cereus VIP3A(a)
34	35	11.7	2444	19	AAV16190	DNA for maize opti
35	33.8	11.3	3157	23	ABL13761	Drosophila melanog
36	33.8	11.3	17330	23	ABL13760	Drosophila melanog
37	33.4	11.1	5903	25	ABZ23939	Nucleotide sequenc
38	32.6	10.9	39061	24	ABS78981	E. coli CFT073 gen
39	32.6	10.9	39061	24	ABS78985	E. coli CFT073 gen
40	32.4	10.8	5892	22	AAS00022	Human cDNA encodin
41	32.4	10.8	6754	25	ABZ34760	Coding sequence SE
42	32.4	10.8	7080	22	ABA08953	Human plexin-B1/SE
43	32.2	10.7	2546	21	AAA09155	Human BMDSP-2 codi
44	31.8	10.6	468	22	ABA42112	Human breast cell
45	31.8	10.6	468	22	ABA52534	Human foetal liver

#### ALIGNMENTS

RESULT 1  
ABQ78752  
ID ABQ78752 standard; cDNA; 300 BP.  
XX AC ABQ78752;  
XX DT 05-DEC-2002 (first entry)  
XX DE Differentially expressed gene RTA00000596F.d.12.1.  
XX KW Differentially expressed gene; cancer; breast cancer; colon cancer;  
KW lung cancer; prostate cancer; expressed sequence tag; EST;  
KW heat shock factor binding protein; CGI-122 gene; quiescin Q6;  
KW moderately-differentiated endometrial adenocarcinoma; TGIF protein;  
KW treacher collins syndrome protein; annexin IV; cyclophilin C;  
KW MHC class I lymphocyte antigen; HLA-E class I MRNA;  
KW glomerulosclerosis gene; 265 proteasome subunit p55;  
KW gamma interferon inducible protein; cyclin-dependent protein kinase;  
KW mitochondrial dodecenoyl-CoA delta-isomerase;  
KW serine hydroxymethyltransferase;  
KW DNA-damage-inducible RNA binding protein; ss.  
XX OS Homo sapiens.

XX US2002076735-A1.  
XX 20-JUN-2002.  
XX 10-MAY-2001; 2001US-0854124.  
XX 25-SEP-1998; 98US-101900P.



CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AB92446 to  
 CC AB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 858 BP; 227 A; 214 C; 222 G; 195 T; 0 other;

Query Match 81.5%; Score 244.4; DB 22; Length 858;  
 Best Local Similarity 95.1%; Pred. No. 1.4e-68;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;  
 QY 1 ATCATGCTTCAGACACATCCCGAGGCGAGACGAAATCCGACCTGGTCAAGGATATGT 60  
 Db 338 ATCATGCTTCAGACACATCCCGAGGCGAGACGAAATCCGACCTGGTCAAGGATATGT 397  
 QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120  
 Db 398 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 457  
 QY 121 CACATGCCAAGCTGATTAAGCTTGTGATGGAGATCAACACGCGGGACTTTCCTCA 180  
 Db 458 CACATGCCAAGCTGATTAAGCTTGTGATGGAGATCAACACGCGGGACTTTCCTCA 517  
 QY 181 CACAAGCGCTCAACACATGTACAACTCCGACGACCACTCCAGCTCTGGAAGACCTC 240  
 Db 518 CACAAGCGCTCAACACATGTACAACTCCGACGACCACTCCAGCTCTGGAAGAGTACTC 577  
 QY 241 AG--CTAGGACTTCTA--AAAAAGGCGTGGTGGAGCGCTTGGTGGG 285  
 Db 578 AGTCTCAGGACTCTAGAGAAAGCGCTGGTGGAGCGGCTTCTGGG 625

## RESULT 3

ACA54679  
 ID ACA54679 standard; cDNA; 1171 BP.

ACA54679;

05-JUN-2003 (first entry)

Human NF-kappaB associated polynucleotide sequence #119.

Human; nuclear factor-kappaB; NF-kappaB; immune disorder; cancer;  
 inflammatory disorder; apoptosis; hepatic disorder; Hodgkin's lymphoma;  
 haematopoietic tumour; hyper-IGM syndrome; viral infection; asthma;  
 hypohidrotic ectodermal dysplasia; human immunodeficiency virus; HIV;  
 X-linked anhidrotic ectodermal dysplasia; al incontinentia pigmenti;  
 influenza; rheumatoid arthritis; inflammatory bowel disease; colitis;  
 atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 experimental allergic encephalomyelitis; autoimmune disorder; wound;  
 hyper immune activity; acute phase response; hypercongenital condition;  
 birth defect; necrotic lesion; organ transplant rejection; pancreas;  
 signal transduction; hyperproliferative disorder; diabetes mellitus;  
 vitamin B12 malabsorption; neurological disorder; Huntington's chorea;  
 Turner's syndrome; bacterial infection; cardiovascular disorder;  
 infertility; psoriasis; haemolytic anaemia; antinflammatory; anti-HIV;  
 cytototoxic; hepatotropic; virucide; antirheumatic; antiarthritic;  
 antiasthmatic; immunomodulator; antidiabetic; antibacterial;  
 neuroprotective; immunosuppressive; vulnery; antibacterial;  
 antinfertility; antianemic; antipsoriatic; cerebroprotective;  
 cardiant; antiarteriosclerotic; gene; ss.

Homo sapiens.

WO200286076-A2.

31-OCT-2002.

19-APR-2002; 2002WO-US12636.

19-APR-2001; 2001US-284962P.

26-APR-2001; 2001US-286645P.

PR

XX

PA

XX

PI

XX

DR

XX

DR

XX

PT

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match 81.5%; Score 244.4; DB 25; Length 1171;  
 Best Local Similarity 95.1%; Pred. No. 1.6e-68;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACACATCCCGAGGCGAGACGAAATCCGACCTGGTCAAGGATATGT 60

Db 378 ATCATGCTTCAGACACATCCCGAGGCGAGACGAAATCCGACCTGGTCAAGGATATGT 437

QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 120

Db 438 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGG 497

QY 121 CACATGCCAAGCTGATTAAGCTTGTGATGGAGATCAACACGCGGGACTTTCCTCA 180

Db 498 CACATGCCAAGCTGATTAAGCTTGTGATGGAGATCAACACGCGGGACTTTCCTCA 557

QY 181 CACAAGCGCTCAACACATGTACAACTCCGACGACCACTCCAGCTCTGGAAGACCTC 240

Db 558 CACAAGCGCTCAACACATGTACAACTCCGACGACCACTCCAGCTCTGGAAGAGTACTC 617

QY 241 AG--CTAGGACTTCTA--AAAAAGGCGTGGTGGAGCGGCTTGGTGGG 285

Db 618 AGTCTCAGGACTTCTAGAGAAAGGCGTGGTGGAGCGGCTTGGTGGG 665

## RESULT 4

AAC77820

ID AAC77820 standard; cDNA; 1210 BP.

XX AAC77820;

AC

XX

08-FEB-2001 (first entry)  
Human cancer associated gene sequence SEQ ID NO:214.  
Human: cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnarary; immunomodulator; antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral; antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant; dermatologic; neuroprotective; thrombolytic; coagulant; nootropic; vasotrophic; antiproliferative; angiogenic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening; ss.  
Homo sapiens.  
WO200053350-A1.  
21-SEP-2000.  
08-MAR-2000; 2000WO-US05882.  
12-MAR-1999; 99US-0124270.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Ruben SM;  
WPI: 2000-587533/55.  
P-FSDB; AAB43611.  
Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer -  
Claim 1; Page 785-786; 2352pp; English.  
AAC77607 to AAC78448 encode the human cancer associated proteins given in AAB43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; vulnarary; immunomodulator; antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral; antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant; dermatologic; neuroprotective; thrombolytic; coagulant; nootropic; vasotrophic; antiproliferative; angiogenic. The polynucleotides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.  
Sequence 1210 BP; 320 A; 300 C; 320 G; 268 T; 2 other;

Query Match 81.5%; Score 244.4; DB 21; Length 1210;  
Best Local Similarity 95.1%; Pred. No. 1.6e-68;  
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 ATCATGCTTCAGACAACATCCGAGGACGACGAAATCCGACCTGGTCAAGATATGT 60  
DB 382 ATCATGCTTCAGACAACATCCGAGGACGACGAAATCCGACCTGGTCAAGATATGT 441  
QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGTTGTGAGACAGCAGGAGG 120  
DB 442 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGTTGTGAGACAGCAGGAGG 501

QY 121 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACACGCGGACCTTCTCTCA 180  
DB 502 CACATGCCAAGCTGGATTAACCTTGACCTTGATGGAGATCAACACACGCGGACCTTCTCTCA 561  
QY 181 CACAAGCGCTCAACACATGATGACAACTCCGACGAACTCCAGCCTCTCGAAGACCTC 240  
DB 562 CACAAGCGCTCAACACATGATGACAACTCCGACGAACTCCAGCCTCTCGAAGACCTC 621  
QY 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCCCTGGTGGGG 285  
DB 622 AGTCTCAGGACTTCTAGAGAAAGCGCTGGTGCAGCCGCTGCTGGGG 669

RESULT 5  
AAH05624  
ID AAH05624 standard; cDNA; 667 BP.  
AC AAH05624;  
XX 26-JUN-2001 (first entry)  
DE Human cDNA clone (5'-primer) SEQ ID NO:2459.  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
KW Homo sapiens.  
OS  
PN EP1074617-A2.  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000EP-0116126.  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX (HELI-) HELIX RES INST.  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI: 2001-318749/34.  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX Claim 1; SEQ ID 2459; 2537pp + CD ROM; English.  
XX The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632





```

Db      241 GCAGGCGGCTGCTGGGG 258
||||| ||| | | |||||
RESULT 7
AAS87103
ID AAS87103 standard; cDNA; 678 BP.
XX
AC AAS87103;
XX
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #22907.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
DR P-PSDB; ABG22916.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
Claim 1; SEQ ID No 22907; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human
diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
Sequence 678 BP; 177 A; 185 C; 206 G; 110 T; 0 other;
XX
Query Match 68.58; Score 205.4; DB 23; Length 678;
Best Local Similarity 91.08; Pred. No. 5e-56;
Matches 252; Conservative 0; Mismatches 21; Indels 4; Gaps 3;
Oy 12 GACAACTATCCGAGGAGGAGCAAGAAATCCGGACCTGCTCAAGGATATGTTGGACACTCGT 71
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 GACAACTATCCGAGGAGGAGCAAGAAATCCGGACCTGCTCAAGGATATGTTGGACACTCGT 441

```

```

Oy 72 ATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAG 131
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 ATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGAGGCAACAGAGGCACATGCCAAG 501
Oy 132 CTGGATAACTTGACCTTGATGGAGATCAACACCAGCGGGACTTTTCCTCACACAAGCGCTC 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 CCGGATAACTTGACCTTGATGGAGATCAACCCCAAGCGGGACTTTTCCTCACACAAGCGCTC 561
Oy 192 AACCATGTGTACAAACTCCGACGAACTCCAGGCTCTGGAAGACCTCAG--CTAGGAC 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 AACCATGTGTACAAACT--CGCAACAACTCCAGGCTTCGGAGAGTACTCAGTCTCAGGAC 620
Oy 250 TTCTA-AAAAAGGCTGTCAGCGCTTGGTTGGGG 285
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 621 TTCTAGAGAAAGGCTGTCAGCGCTTGGTTGGGG 657

```

```

RESULT 8
AAS87105
ID AAS87105 standard; cDNA; 678 BP.
XX
AC AAS87105;
XX
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #22909.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
DR P-PSDB; ABG22916.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
Claim 1; SEQ ID No 22909; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human

```

CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 678 BP; 177 A; 185 C; 206 G; 110 T; 0 other;

Query Match 68.5%; Score 205.4; DB 23; Length 678;  
Best Local Similarity 91.0%; Pred. No. 5e-56;  
Matches 252; Conservative 0; Mismatches 21; Indels 4; Gaps 3;

QY 12 GACACATCCCAAGGCAGACGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGT 71  
|||||  
Db 382 GACAACATCCCAAGGCAGACGAAATCCGGACCCCTGGTCAAGGATATATGGACACTCGT 441  
|||||

QY 72 ATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAG 131  
|||||  
Db 442 ATAGCCAAACTCCGAGTGTCTGTGACAGCTTTGTGAGGCAACAGAGGCACATGCCAAG 501  
|||||

QY 132 CTGGATACTTACCTTGATGGAGATCAACACAGCGGGGACTTTCCTCACAAAGCGCTC 191  
|||||  
Db 502 CGGGATACTTACCTTGATGGAGATCAACACAGCGGGGACTTTCCTCACAAAGCGCTC 561  
|||||

QY 192 AACCATCTACAACTCCGACGACCTCCAGCCTCTGGAAGACCTCAG--CTAGGAC 249  
|||||  
Db 562 AACCATCTACAACT--CGCACAACTCCAGCCTCTCGGAGAGTACTCAGTCTCAGGAC 620  
|||||

QY 250 TTCTA-AAAAAGCCTGGTGCAGCGCCTTGGTTGGGG 285  
|||||  
Db 621 TTCTAGAAAGCCTGGTGCAGGACGCTTGTCTGGGG 657  
|||||

RESULT 9  
AAS87106  
ID AAS87106 standard; cDNA; 1190 BP.  
AC AAS87106;  
XX  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX DNA encoding novel human diagnostic protein #22910.  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX P-PSDB; ABG22919.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity -  
XX  
XX Claim 1; SEQ ID No 22910; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1190 BP; 298 A; 304 C; 310 G; 278 T; 0 other;

Query Match 64.7%; Score 194; DB 23; Length 1190;  
Best Local Similarity 92.5%; Pred. No. 3e-52;  
Matches 270; Conservative 0; Mismatches 15; Indels 7; Gaps 6;

QY 1 ATCATGCTTCAGAACATCCCGAAGCGAGACGAAATCCGGACCCCTGGTCAAGGATATGT 60  
|||||  
Db 383 ATCATGCTTCAGAACATCCCGAAGCGAGACGAAATCCGGACCCCTGGTCAAGGATATGT 442  
|||||

QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGC--TTTGTGAGACAGCAGGAG 119  
|||||  
Db 443 GGGACACTCGTATAGCCAACTCCGAGTGTCTGTGACAGC--TTTGTGAGACAGCAGGAG 502  
|||||

QY 120 GCACATGCCCAAGCT--GGATAACTTGACCTT-GATGGAGATCAACACCGCGG-GACTTTC 176  
|||||  
Db 503 GCACATGCCCAAGCTCGGATTAACCTTGACTTTCGATGGAGATCCACACCGCGGTGACTTTC 562  
|||||

QY 177 CTCACACAAAGCGCTCAACCCACATGTACAACTCCGCACGAACTCCAGCCTCTCGGAAAGA 236  
|||||  
Db 563 CTCACACAAAGCGCTCAACCCACATGTACAACTCCGCACGAACTCCAGCCTCTCGGAGAGT 622  
|||||

QY 237 CCTCAG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTGGGG 285  
|||||  
Db 623 ACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGTGCAGCGCGCTTGTCTGGGG 674  
|||||

RESULT 10  
AAS87104  
ID AAS87104 standard; cDNA; 1499 BP.  
AC AAS87104;  
XX  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX DNA encoding novel human diagnostic protein #22908.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;



CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 454 BP; 128 A; 113 C; 116 G; 97 T; 0 other;

Query Match 52.7%; Score 158; DB 24; Length 454;  
Best Local Similarity 98.9%; Pred. No. 8.3e-41;  
Matches 180; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
QY 1 ATCATGCTTCACACATCCCGAAGCGCAGACGAATCCGGACCTGGTCAAGGATATGT 60  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
274 ATCATGCTTCACACATCCCGAAGCGCAGACGAATCCGGACCTGGTCAAGGATATGT 332  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACAGCAGGAGG 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
333 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACAGCAGGAGG 392  
QY 121 CACATGCCAAGCTGGATACCTTGACCTTGATGGAGATCAACACCGCGGACTTTCCCTCA 180  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
393 CACATG-CAAGCTGGATACCTTGACCTTGATGGAGATCAACACCGCGGACTTTCCCTCA 451  
QY 181 CA 182  
DB ||  
452 CA 453

RESULT 12  
ABQ59992  
ID ABQ59992 standard; cDNA; 603 BP.  
AC ABQ59992;  
XX  
DT 02-AUG-2002 (first entry)  
DE Human colon cancer related nucleotide sequence SEQ ID NO:3687.  
XX  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX  
XX Homo sapiens.  
PN WO200229086-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001; 2001WO-US30732.  
PF  
XX 02-OCT-2000; 2000US-237271P.  
PR  
XX (FARB ) BAYER CORP.  
PA  
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
PI Thiagalingam A, Lewis ME;  
XX  
XX WPI; 2002-426115/45.  
DR  
XX New isolated nucleic acid that is differentially expressed in cancer  
XX tissues useful for determining the presence of colon cancer in a cell  
XX or tissue type, and in antisense therapy -  
XX  
XX Claim 1; Fig 1; 796pp; English.  
XX  
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
XX expressed in cancer tissues. ABB78993 to ABB79004 represent proteins  
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be

CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists.  
XX  
SQ Sequence 603 BP; 157 A; 161 C; 162 G; 117 T; 6 other;

Query Match 43.5%; Score 130.6; DB 24; Length 603;  
Best Local Similarity 90.6%; Pred. No. 6.5e-32;  
Matches 184; Conservative 0; Mismatches 11; Indels 8; Gaps 4;  
QY 1 ATCATGCTTCACACATCCCGAAGCGCAGACGAATCCGGACCTGGTCAAGGATATGT 60  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
364 ATCATGCTTCACACATCCCGAAGCGCAGACGAATCCGGACCTGGTCAAGGATATGT 423  
QY 61 GGGACACTCGTATAGCCAACTCCGAGTGTCTGCTGACAGCTTTGTGACAGCAGGAGG 120  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
424 GGGACACTCGTATAGCCAACTCCGA-CGTCTGCTGACAGCTTTGTGACAGCAGGAGG 482  
QY 121 CACATGCCAAGCTGGATACCTTGACCTTGATGGAGATCAACACCGCGGACTTT 173  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
483 CACATGCCAAGCTGGATACCTTGACCTTGATGGAGATCAACACCGCGGACTTT 542  
QY 174 TTCCTCACACAAAGCGCTCAACCA 196  
DB |||||||||||||||||||||||||||  
543 TCCTTCACACAGCGCTNAACCA 565

RESULT 13  
ABQ59972  
ID ABQ59972 standard; cDNA; 649 BP.  
AC ABQ59972;  
XX  
DT 02-AUG-2002 (first entry)  
DE Human colon cancer related nucleotide sequence SEQ ID NO:3667.  
XX  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX  
XX Homo sapiens.  
PN WO200229086-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001; 2001WO-US30732.  
PF  
XX 02-OCT-2000; 2000US-237271P.  
PR  
XX (FARB ) BAYER CORP.  
PA  
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
PI Thiagalingam A, Lewis ME;  
XX  
XX WPI; 2002-426115/45.  
DR  
XX New isolated nucleic acid that is differentially expressed in cancer  
XX tissues useful for determining the presence of colon cancer in a cell  
XX or tissue type, and in antisense therapy -  
XX  
XX Claim 1; Fig 1; 796pp; English.  
XX





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:00:56 ; Search time 1477 Seconds  
(without alignments)  
8309.335 Million cell updates/sec

Title: US-09-854-124-7  
Perfect score: 300  
Sequence: 1 atcagcttcagacaacatc.....tggggattaacccttcagac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	244.4	81.5	858	6	BD156103	BD156103 Primer fo
2	244.4	81.5	858	9	AK001275	AK001275 Homo sapi
3	244.4	81.5	1174	9	AF151880	AF151880 Homo sapi
4	244.4	81.5	1180	9	BC010164	BC010164 Homo sapi
5	244.4	81.5	1188	9	AF201939	AF201939 Homo sapi
6	243	81.0	667	6	BD147616	BD147616 Primer fo
7	242.8	80.9	1203	9	BC003186	BC003186 Homo sapi
8	216.4	72.1	116792	9	HS1100H13	AL035419 Human DNA
9	158	52.7	454	6	AX336730	AX336730 Sequence
10	136.4	45.5	933	5	AB097169	AB097169 xenopus 1
11	128.2	42.7	160356	9	AC009068	AC009068 Homo sapi
12	128.2	42.7	185321	9	AC123908	AC123908 Homo sapi
13	128.2	42.7	245210	2	AC137771	AC137771 Homo sapi
14	126.6	42.2	3097	9	AK091519	AK091519 Homo sapi
15	99.6	33.2	188292	2	AC103360	AC103360 Mus muscu
16	94	31.3	207307	2	AC117085	AC117085 Rattus no
17	59.6	19.9	825	3	AY071234	AY071234 Drosophill
18	59.6	19.9	72722	3	AC004373	AC004373 Drosophill
19	59.6	19.9	132666	3	AC005270	AC005270 Drosophill
20	59.6	19.9	182897	3	AC092232	AC092232 Drosophill
21	59.6	19.9	218794	2	AC018064	AC018064 Drosophill
22	59.6	19.9	292393	3	AE003576	AE003576 Drosophill
23	47.2	15.7	213	9	AF534827	AF534827 Homo sapi
24	47.2	15.7	92814	9	AL391903	AL391903 Human DNA
25	47.2	15.7	219443	2	AC023320	AC023320 Homo sapi
26	44.6	14.9	165158	2	AC141687	AC141687 Apis mell
27	40	13.3	1150	9	AF534824	AF534824 Homo sapi
28	40	13.3	91511	9	AL133404	AL133404 Human DNA
29	39.6	13.2	734	6	BD150236	BD150236 Primer fo
30	39.6	13.2	3071	6	BD158620	BD158620 Primer fo
31	39.6	13.2	3071	9	AK023755	AK023755 Homo sapi
32	36.2	12.1	225145	2	AC111766	AC111766 Rattus no
33	36.2	12.1	241188	2	AC133023	AC133023 Rattus no
34	36.2	12.0	654	6	AX386100	AX386100 Sequence
35	36	12.0	49995	2	AC011335	AC011335 Homo sapi
36	36	12.0	77139	2	AC008466	AC008466 Homo sapi
37	36	12.0	120849	2	AC087148	AC087148 Mus muscu
38	36	12.0	128928	9	AC068658	AC068658 Homo sapi
39	36	12.0	157208	9	AC113409	AC113409 Homo sapi
40	36	12.0	203418	2	AC063958	AC063958 Homo sapi
41	35.8	11.9	1935	6	AX078491	AX078491 Sequence
42	35.8	11.9	2746	8	NCACU9	X56672 N.crassa ma
43	35.6	11.9	1129	8	AF499724	AF499724 Theillungi
44	35.6	11.9	215707	2	AC099096	AC099096 Rattus no
45	35.6	11.9	215707	2	AC099096	AC099096 Rattus no

ALIGNMENTS

RESULT 1	BD156103	858 bp	DNA	linear	PAT 17-JAN-2003
LOCUS	BD156103				
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD156103				
VERSION	BD156103.1	GI:27861861			
KEYWORDS	JP 2002191363-A/10946.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 858)				
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,				
	Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				

Patent: JP 2002191363-A 10946 09-JUL-2002;  
**JOURNAL**  
**HELIIX RESEARCH INSTITUTE**  
**COMMENT**  
 PN JP 2002191363-A/10946  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 PI KETICHI NAGAI, TETSUJI OTSUKI  
 PC  
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH key  
 Location/Qualifiers  
 FT CDS  
 (37)..(591).  
 FEATURES  
 source  
 1..858  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 227 a 214 c 222 g 195 t  
 ORIGIN  
 Query Match 81.5%; Score 244.4; DB 6; Length 858;  
 Best Local Similarity 95.1%; Pred. No. 9.6e-55;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;  
 QY 1 ATCATGCTTCAGACAACTCCGGAAGCGACGAGCAAAATCCGGACCTGGTCAAGGATATGT 60  
 DB 338 ATCATGCTTCAGACAACTCCGGAAGCGACGAGCAAAATCCGGACCTGGTCAAGGATATGT 397  
 QY 61 GGGACACTCGTATAGCAAACTCCGGAAGCGCTGCTGCTGACAGCTTTGTGAGACACGAGG 120  
 DB 398 GGGACACTCGTATAGCAAACTCCGGAAGCGCTGCTGCTGACAGCTTTGTGAGACACGAGG 457  
 QY 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACGAGCGGACTTTCCTCA 180  
 DB 458 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACGAGCGGACTTTCCTCA 517  
 QY 181 CACAGCGCTCAACACATGTACAACTCCGACAGCACTCCAGCCCTCTGGAAGACCCTC 240  
 DB 518 CACAGCGCTCAACACATGTACAACTCCGACAGCACTCCAGCCCTCTGGAAGACTC 577  
 QY 241 AG--CTAGGACTTCTA-AAAAAGCGCTGGTGCAGCGCTTGGTGGG 285  
 DB 578 AGCTTCAGGACTTCTAGAGAAGCGCTGGTGCAGCGCGCTTGGTGGG 625  
 RESULT 2  
 AK001275 858 bp mRNA linear PRI 01-AUG-2002  
 LOCUS  
 Homo sapiens cDNA FLJ10413 fis, clone NT2RP1000063.  
 AK001275  
 ACCESSION  
 AK001275.1 GI:7022428  
 VERSION  
 oligo capping; fis (full insert sequence).  
 KEYWORDS  
 Homo sapiens (human)  
 SOURCE  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 ISOGAI, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
 Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
 Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
 Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
 Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 858)  
 ISOGAI, T. and OTSUKI, T.  
 TITLE  
 JOURNAL  
 AUTHORS

**TITLE** Identification of novel human genes evolutionarily conserved in  
**JOURNAL** Caenorhabditis elegans by comparative proteomics  
**MEDLINE** Genome Res. 10 (5), 703-713 (2000)  
**PUBMED** 20272150  
**REFERENCE** 2 (bases 1 to 1174)  
**AUTHORS** Lin, W.-C.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia  
 Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan  
**FEATURES** Location/Qualifiers  
 source  
 1..1174  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 77..634  
 /codon\_start=1  
 /product="CGI-122 protein"  
 /protein\_id="AAD34117.1"  
 /db\_xref="GI:4929713"  
 /translation="MDAAEEVEFLAEKELVTIIPNFSLDKIYILIGDLGPFNPGLPVEV  
 PLWLAINLQKQKRLPPEWMDVEKMRDHERKEETFTPMPSPYMELTKLLNH  
 ASDNTPKADEIRTLVKDMDTRIAKLRVSADSFVYRQEAHAKLDNLTLMEINTSGTFL  
 TQALNHYKLRLNLPLESTQSQDF"  
**BASE COUNT** 300 a 299 c 307 g 268 t  
**ORIGIN**  
 Query Match 81.5%; Score 244.4; DB 9; Length 1174;  
 Best Local Similarity 95.1%; Pred. No. 9.2e-55;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;  
 QY 1 ATCATGCTTCAGACAACATCCGGAAGGCAGAGCAAAATCCGGACCCTGGTCAAGGATATGT 60  
 |||||  
 DB 378 ATCATGCTTCAGACAACATCCGGAAGGCAGAGCAAAATCCGGACCCTGGTCAAGGATATGT 437  
 QY 61 GGGACACTCGTATACCAACTCCGAGTGTCTGTGACAGCTTTGTGACAGCAGGAGG 120  
 |||||  
 DB 438 GGGACACTCGTATACCAACTCCGAGTGTCTGTGACAGCTTTGTGACAGCAGGAGG 497  
 QY 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACACCGGAGCTTTCCCTCA 180  
 |||||  
 DB 498 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACACCGGAGCTTTCCCTCA 557  
 QY 181 CACAAGCGCTCAACACATGATCAAACTCCGACGAACTCCAGCTCTCGAAGACCTC 240  
 |||||  
 DB 558 CACAAGCGCTCAACACATGATCAAACTCCGACGAACTCCAGCTCTCGAAGACTACTC 617  
 QY 241 AG--CTAGGACTTCTA-AAAAAGGCTGGTGCAGCGCTTGGTGGG 285  
 || |||||  
 DB 618 AGTCTCAGGACTTCTAGAGAAAGGCTGGTGCAGCGCTTGGTGGG 665  
**RESULT 4**  
**BC010164**  
**LOCUS** BC010164  
**DEFINITION** Homo sapiens, HSPC037 protein, clone MGC:19836 IMAGE:4098007, mRNA,  
 complete cds.  
**ACCESSION** BC010164  
**VERSION** BC010164.1 GI:14603431  
**KEYWORDS** MGC.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 1180)  
**AUTHORS** Strausberg, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (02-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>

## COMMENT

Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,  
 Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 27 Row: 1 Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Genomescan gene  
 prediction, Similarity but not identity to protein.

## FEATURES

Location/Qualifiers  
 1..1180  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="LocusID:51659"  
 /db\_xref="taxon:9606"  
 /clone="MGC:19836 IMAGE:4098007"  
 /tissue\_type="Brain, neuroblastoma"  
 /clone\_lib="NIH\_MGC\_19"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 68..625  
 /codon\_start=1  
 /product="HSPC037 protein"  
 /protein\_id="AAH10164.1"  
 /db\_xref="GI:14603432"  
 /translation="MDAAEEVEFLAEKELVTIIPNFSLDKIYILIGDLGPFNPGLPVEV  
 PLWLAINLQKQKRLPPEWMDVEKMRDHERKEETFTPMPSPYMELTKLLNH  
 ASDNTPKADEIRTLVKDMDTRIAKLRVSADSFVYRQEAHAKLDNLTLMEINTSGTFL  
 TQALNHYKLRLNLPLESTQSQDF"  
**BASE COUNT** 316 a 292 c 307 g 265 t  
**ORIGIN**

Query Match 81.5%; Score 244.4; DB 9; Length 1180;  
 Best Local Similarity 95.1%; Pred. No. 9.2e-55;  
 Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;  
 QY 1 ATCATGCTTCAGACAACATCCGGAAGGCAGAGCAAAATCCGGACCCTGGTCAAGGATATCT 60  
 |||||  
 DB 369 ATCATGCTTCAGACAACATCCGGAAGGCAGAGCAAAATCCGGACCCTGGTCAAGGATATCT 428  
 QY 61 GGGACACTCGTATACCAACTCCGAGTGTCTGTGACAGCTTTGTGACAGCAGGAGG 120  
 |||||  
 DB 429 GGGACACTCGTATACCAACTCCGAGTGTCTGTGACAGCTTTGTGACAGCAGGAGG 488  
 QY 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACCGGAGCTTTCCCTCA 180  
 |||||  
 DB 489 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACCGGAGCTTTCCCTCA 548  
 QY 181 CACAAGCGCTCAACACATGATCAAACTCCGACGAACTCCAGCTCTCGAAGACCTC 240  
 |||||  
 DB 549 CACAAGCGCTCAACACATGATCAAACTCCGACGAACTCCAGCTCTCGAAGACTACTC 608  
 QY 241 AG--CTAGGACTTCTA-AAAAAGGCTGGTGCAGCGCTTGGTGGG 285  
 || |||||  
 DB 609 AGTCTCAGGACTTCTAGAGAAAGGCTGGTGCAGCGCTTGGTGGG 656

## RESULT 5

```

AF201939      1188 bp      mRNA      linear      PRI 20-JUL-2000
LOCUS        Homo sapiens DC5 (DC5) mRNA, complete cds.
DEFINITION   AF201939
ACCESSION    AF201939
VERSION      AF201939.1 GI:9295181
KEYWORDS     FLI CDNA.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.
JOURNAL      Novel genes expressed in human dentritic cell
REFERENCE    2 (bases 1 to 1188)
AUTHORS      Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.
TITLE        Direct Submission
JOURNAL      Submitted (05-NOV-1999) Chinese National Human Genome Center at
              Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
              Shanghai 201203, P. R. China
FEATURES     Location/Qualifiers
              1..1188
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /cell_type="dentritic cell"
              2..1188
               /gene="DC5"
               /product="DC5"
               /protein_id="AAP86875.1"
               /db_xref="GI:9295182"
               /translation="MDAAEVEFLAEKELVTIIPNFSLDKIYLLIGDLPENPGLPVEV
               PLTAIINKLORCLLPPEWMDVEKLEKMRDHERKEETFTPMSPYPMELTKILLNH
               ASDNIPKADIRILVKMDWTRIAKLVSADSFVRQEAHKLNLTLMEINTSGIFL
               TQALNHYKRTNLQPLETSQSDP"
BASE COUNT   314 a 299 c 307 g 268 t
ORIGIN
Query Match      81.5%; Score 244.4; DB 9; Length 1188;
Best Local Similarity 95.18; Pred. No. 9.2e-55;
Matches 274; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

Qy 1 ATCATGCTTCAGACAAACATCCGAAAGCAGACAGCAAAATCCGACCCCTGGTCAAGGATATGT 60
    |||||
Db 378 ATCATGCTTCAGACAAACATCCGAAAGCAGACAGCAAAATCCGACCCCTGGTCAAGGATATGT 437
    |||||

Qy 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGACAGCTTTGTGACAGCAGCAGGAGG 120
    |||||
Db 438 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGACAGCTTTGTGACAGCAGCAGGAGG 497
    |||||

Qy 121 CACATGCCAAGCTGATTAACCTTGACCTTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 180
    |||||
Db 498 CACATGCCAAGCTGATTAACCTTGACCTTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 557
    |||||

Qy 181 CACAAGCGCTCAACCAACATGTACAACTCCGACAGAACTCCAGCCCTCTGGAAGACCTC 240
    |||||
Db 558 CACAAGCGCTCAACCAACATGTACAACTCCGACAGAACTCCAGCCCTCTGGAAGTACTC 617
    |||||

Qy 241 AG--CTAGGACTTCTA-AAAAGGCTTGTGTCAGCCGCTTGGTGGGG 285
    || |||||
Db 618 AGTCTCAGGACTTCTTAGAAGAAAGGCTTGTGTCAGCGGGCTTGTGGGG 665
    || |||||

RESULT 6
LOCUS        BD147616
DEFINITION   BD147616
ACCESSION    BD147616
VERSION      BD147616.1 GI:27853374
KEYWORDS     JP 2002191363-A/2459.

```

```

SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
              Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE        Primer for synthesizing full-length cDNA and use thereof
JOURNAL      Patent: JP 2002191363-A 2459 09-JUL-2002;
              HELIX RESEARCH INSTITUTE
COMMENT      OS Homo sapiens (human)
              PN JP 2002191363-A/2459
              PD 09-JUL-2002
              PF 28-JUL-2000 JP 2000280990
              PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
              PI SAITO,
              PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
              PI KEIICHI NAGAI,TETSUJI OTSUKI
              PC
              C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
              10,
              PC C12P21/02,C12O1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
              Primer for synthesizing full-length cDNA and use thereof FH Key
              FT source 1..667
              FT Location/Qualifiers
              1..667
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
BASE COUNT   171 a 176 c 177 g 137 t 6 others
ORIGIN
Query Match      81.08; Score 243; DB 6; Length 667;
Best Local Similarity 94.48; Pred. No. 2.4e-54;
Matches 272; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

Qy 1 ATCATGCTTCAGACAAACATCCGAAAGCAGACAGCAAAATCCGACCCCTGGTCAAGGATATGT 60
    |||||
Db 337 ATCATGCTTCAGACAAACATCCGAAAGCAGACAGCAAAATCCGACCCCTGGTCAAGGATATGT 396
    |||||

Qy 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTCAGCTTTGTGACAGCAGCAGGAGG 120
    |||||
Db 397 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTCAGCTTTGTGACAGCAGCAGGAGG 456
    |||||

Qy 121 CACATGCCAAGCTGATTAACCTTGACCTTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 180
    |||||
Db 457 CACATGCCAAGCTGATTAACCTTGACCTTTGATGGAGATCAACACAGCGGGACTTTCTCTCA 516
    |||||

Qy 181 CACAAGCGCTCAACCAACATGTACAACTCCGACAGAACTCCAGCCCTCTGGAAGACCTC 240
    |||||
Db 517 CACAAGCGCTCAACCAACATGTACAACTCCGACAGAACTCCAGCCCTCTGGAAGTACTC 576
    |||||

Qy 241 AG--CTAGGACTTCTA-AAAAGGCTTGTGTCAGCCGCTTGGTGGGG 285
    || |||||
Db 577 AGTCTCAGGACTTCTTANANAAAGGCTTGTGTCANGCGGCTTGTGGGG 624
    || |||||

RESULT 7
LOCUS        BC003186
DEFINITION   BC003186
ACCESSION    BC003186
VERSION      BC003186.1 GI:13112024
KEYWORDS     MGC.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      1 (bases 1 to 1203)
              Strausberg,R.

```

**TITLE**  
JOURNAL

**REMARK**  
**COMMENT**

Direct Submission  
Submitted (13-FEB-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps@mail.nih.gov](mailto:cgaps@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ruben Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@cgsc.bc.ca](mailto:info@cgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 6 Row: d Column: 16  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4929712.  
**FEATURES**  
source  
1..1203  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:51659"  
/db\_xref="taxon:9606"  
/clone="MGC:673 IMAGE:3503268"  
/tissue\_type="Brain, neuroblastoma"  
/clone\_lib="NIH\_MGC\_19"  
/lab\_host="DH10B-R"  
/note="Vector: pON7"  
85..642  
/codon\_start=1  
/product="HSPC037 protein"  
/protein\_id="AAH03186.1"  
/db\_xref="GI:13112025"  
/translation="MDAAEVEFLAEKELVTIIPNLSDKIYLIIGDPLGPNPLPVEV  
PWLAINKQKQKRLLEPEWMDVEKLEMRDHRKEETFTPMPSPYMELTKLLNH  
ASDNTPKADEITLVKDMMDTRIAKLRVSADSFVRQEAHAKLDNLTLMEINTSGTL  
TQALNHMYKLRTNLOPLESTQSQDF"  
BASE COUNT 326 a 300 c 309 g 268 t  
ORIGIN

Query Match 80.9%; Score 242.8; DB 9; Length 1203;  
Best Local Similarity 94.8%; Pred. No. 2.5e-54;  
Matches 273; Conservative 0; Mismatches 12; Indels 3; Gaps 2;  
QY 1 ATCATGCTTCAGACATCCCGAAGGCGAGAGCAATCCGGACCCCTGGTCAAGGATATGT 60  
|||||  
DB 386 ATCATGCTTCAGACATCCCGAAGGCGAGAGCAATCCGGACCCCTGGTCAAGGATATGT 445  
|||||  
QY 61 GGGACACTCGTATAGCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACAGACAGGAGG 120  
|||||  
DB 446 GGGACACTCGTATAGCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACAGACAGGAGG 505  
|||||  
QY 121 CACATCCCAAGCTGGTAATCTTACCTTGATGGAGATCAACACCGCGGACTTTCCTCA 180  
|||||  
DB 506 CACATCCCAAGCTGGTAATCTTACCTTGATGGAGATCAACACCGCGGACTTTCCTCA 565  
|||||  
QY 181 CACAAGCGCTCAACCATGATCAAACTCCGACAGAACTCCAGCCTCTGGAAAGACCTC 240  
|||||  
DB 566 CACAAGCGCTCAACCATGATCAAACTCCGACAGAACTCCAGCCTCTGGAGAGTACTC 625  
|||||  
QY 241 AG--CTAGACATCTCA-AAAAGGCTGTGGTCAGCGGCTTGGTTGGG 285  
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

**Db** 626 AGTCTCAGGACTTCTAGAGAAAGGCGCTGTGTCAGCGGCTTGCTGGG 673

**RESULT** 8

**LOCUS** HS1100H13

**DEFINITION** Human DNA sequence from clone RP5-1100H13 on chromosome 20q11.2  
Contains the 3' end of gene KIAA1219, a putative novel gene, a DC5  
pseudogene, the gene for a putative RhoGAP domain containing  
protein, ESTs, STSs, GSSs and two CpG islands, complete sequence.

**ACCESSION** AL035419

**VERSION** AL035419.12 GI:11546045

**KEYWORDS** HTG: CpG island; DC5; KIAA1219; RhoGAP.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** 1 (bases 1 to 116792)

**AUTHORS** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**TITLE** Direct Submission

**JOURNAL** Submitted (12-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)

**COMMENT** On Dec 5, 2000 this sequence version replaced gi:10198625.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence  
was generated from part of bacterial clone contigs of human  
Chromosome 20, constructed by the Sanger Centre Chromosome 20  
Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. RP5-1100H13 is from  
the library RPCI-5 constructed by the group of Pieter de Jong. For  
further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: pCYPAC2  
This sequence is the entire insert of clone RP5-1100H13 The true  
left end of clone RP11-12201 is at 47827 in this sequence. The true  
right end of clone RP5-927M24 is at 16404 in this sequence.

**FEATURES**  
Location/Qualifiers  
1..116792  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="RZPD:RPCIP704H131100"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="q11.2"  
/clone="RP5-1100H13"  
/clone\_lib="RPCI-5"  
1331..1465  
/note="MER58C repeat: matches 68..89 of consensus"  
1587..13349  
/gene="djl100H13.1"  
Join(<1587..1723,4379..4487,5260..5338,8641..8789,  
9265..13349)  
/gene="djl100H13.1"  
/product="djl100H13.1.1 (KIAA1219 (similar to Drosophila  
H909358 and C. elegans D2085.5) isoform 1)"  
/note="match: cDNAs: Em:AB033045 Em:AL132998

**repeat\_region**

**gene**

**mRNA**

```

match: ESTs: Em: A1190579 Em: AA256927 Em: N36561 Em: AA874351
Em: A1281190 Em: AAL180094 Em: AA669533 Em: AA511971
Em: A1195307 Em: AA215757 Em: AA156777 Em: AA131113
Em: AA418776 Em: A1357456 Em: A1379408 Em: N76170 Em: W52362
Em: AA024901 Em: A1334294 Em: AA205264 Em: AA215603
Em: AA418686 Em: AA024902 Em: A1379162 Em: AA424825
Em: A1202970 Em: A1248059 Em: AA648752 Em: A1359832
Em: A1366165 Em: A1224862 Em: A1359759 Em: A1492050 Em: R76319
Em: A1050034 Em: A1423927 Em: AA024896 Em: R78702 Em: A1066457
Em: R72124 Em: A1361138 Em: R81299 Em: AA282907 Em: L44403
Em: A1500502 Em: Z98499 Em: R83793 Em: AA412128 Em: AA179405
Em: AA459722 Em: AA173335 Em: AA283091 Em: H26506 Em: H02652
Em: AA258043 Em: N64761 Em: AA255436 Em: A1271673 Em: A1017681
Em: AA459863 Em: A1221344 Em: R65916 Em: W07865 Em: AA383163
Em: R26520 Em: AA872554 Em: AA343687 Em: AA889796 Em: R68642
Em: AA580081 Em: AA156671 Em: AA552080 Em: AA024895"
/evidence-not_experimental
join(<1587, 1723, 4379, 4487, 5260, 5338, 8641, 8789,
9265, 9458)
/gene="dJ1100H13.1"
/note="Continues in Em: AL049868 as dJ927M24.2
match: proteins: Tr: Q9Y3G9 Tr: Q9W2X2 Tr: Q9U4G6 Tr: Q18431"
/codon_start=3
/evidence-not_experimental
/product="dJ1100H13.1.1 (K1AA1219 (similar to Drosophila
GH03358 and C. elegans D2085.5) isoform 1)"
/protein_id="CAB41646.1"
/db_xref="GI: 4704208"
/db_xref="SPTREMBL: Q9Y3G9"
/translation="DSLESNTSDSDSNMDLMPILKOPSLTLELPNHTDNLSSQ
RLSPSSRMRLPOCRPVPLPGPPTRVSVWVDDINPEPLSMTEISGVEYTTAN
SSTSRLTLEKEVPFIHPLNLTGLFRKIQTGATGFNMVPLVDGMI VSRALGEL
VRQVVICRRKRLESDSYSPHVRKQKITDIYKNRKQLEPFTYLSLQFVGLKN
CSS"
repeat_region 2753..3187
/note="L1M4 repeat: matches 3867. 4299 of consensus"
4727..5189
/gene="dJ1100H13.1"
/note="match: GSS: Em: A0663797"
4732..4875
/gene="dJ1100H13.1"
/note="match: STS: Em: HSPF04E1"
4745..5581
/gene="dJ1100H13.1"
/note="match: GSS: Em: A0745415"
complement(4997..5442)
/note="match: STS: Em: G33216"
complement(5210..5903)
/note="match: GSS: Em: A0321022"
5522..6464
/note="L1M40 repeat: matches 5388. 6322 of consensus"
6837..6970
/note="MIR repeat: matches 32. 175 of consensus"
complement(6984..7365)
/note="match: GSS: Em: A0807920"
7232..7391
/note="L1M4 repeat: matches 7614. 7783 of consensus"
7610..7684
/note="AluJb repeat: matches 39. 113 of consensus"
join(<8150, 8789, 9265, 9431, 10364, >10484)
/gene="dJ1100H13.1"
/product="dJ1100H13.1.2 (K1AA1219, isoform 2)"
/note="match: ESTs: Em: AL536595"
/evidence-not_experimental
8567..8616
/note="2 copies 25 mer 98% conserved"
12187..12234
/note="12 copies 4 mer caca 89% conserved"
12536..12571
/note="18 copies 2 mer aa 80% conserved"
12911..12955
/note="L2 repeat: matches 2657. 2701 of consensus"
complement(12943..13346)

```

```

polyA_signal /note="match: STS: Em: G24573"
13322..13327
/gene="dJ1100H13.1"
13349
/gene="dJ1100H13.1"
14509..14637
/note="MERSA repeat: matches 54. 182 of consensus"
15718..22954
/gene="dJ1100H13.2"
join(<15718, 15865, 19926, 20186)
/gene="dJ1100H13.2"
/product="dJ1100H13.2.2 (putative novel protein (isoform
2))"
/note="match: ESTs: Em: AA776833 Em: AA972215"
/evidence-not_experimental
join(<15797, 15865, 20520, 20634, 22592, 22954)
/gene="dJ1100H13.2"
/product="dJ1100H13.2.1 (putative novel protein (isoform
1))"
/note="match: ESTs: Em: A190489 Em: A1656415 Em: AA974242"
/evidence-not_experimental
complement(15882..16370)
/note="match: STS: Em: HS927M24S"
16047..16178
/note="4 copies 33 mer 91% conserved"
16227..16339
/note="MER91A repeat: matches 2. 115 of consensus"
16620..16889
/note="AluJb repeat: matches 2. 289 of consensus"
17303..17608
/note="Alusq repeat: matches 1. 306 of consensus"
17634..17809
/note="MIR repeat: matches 79. 262 of consensus"
17825..18011
/note="MIR repeat: matches 36. 238 of consensus"
19117..19233
/note="MIR repeat: matches 126. 252 of consensus"
19275..19565
/note="AlusC repeat: matches 1. 292 of consensus"
19584..19698
/note="MIR repeat: matches 132. 261 of consensus"
20186
/gene="dJ1100H13.2"
join(<20613, 20748, 22592, 22954)
/gene="dJ1100H13.2"
/product="dJ1100H13.2.3 (putative novel protein (isoform
3))"
/note="match: ESTs: Em: A1632573"
/evidence-not_experimental
20940..21205
/note="L2 repeat: matches 2403. 2748 of consensus"
21246..21354
/note="MIR repeat: matches 127. 244 of consensus"
21421..21616
/note="L2 repeat: matches 1610. 1811 of consensus"
21977..22107
/note="MER45C repeat: matches 1. 907 of consensus"
22215..22528
/note="AlusX repeat: matches 1. 311 of consensus"
72.1%; Score 216.4; DB 9; Length 116792;
Best Local Similarity 91.3%; Pred. No. 1.6e-47;
Matches 263; Conservative 0; Mismatches 21; Indels 4; Gaps 3;
QY 1 ATCATGCTTCAGACAAACATCCGAGGACGACGAAATCCGACCTGGTCAAGATATGT 60
|||||
Db 49659 ATCATGCTTCAGACAAACATCCGAGGACGACGAAATCCGACCTGGTCAAGATATAT 49600
QY 61 GGGACACTCGTATAGCCAAACTCGAGTGTCTGTGACAGTTTGTGACAGACGACGAGG 120
|||||
Db 49599 GGGACACTCGTATAGCCAAACTCGAGTGTCTGTGACAGTTTGTGACGACGAGG 49540
QY 121 CACATGCCAAGCTGGATAACTTGACCTTGATGGAGATCAACACACGCGGACTTTCCTCA 180

```

```

||||| 49539 CACATGCCAAGCCGGATACCTTACCTTGATGGATCAACCCCGGGGACTTTCCTCA 49480
||||| 181 CACAAGCGCTCAACACATGTACAACTCCGACCAACCTCCAGCCTCTGGAAAGACCTC 240
||||| 49479 CACAAGCGCTCAACACATGTACAACTCCGACCAACCTCCAGCCTCTGGAGAGTACTC 49421
||||| 241 AG--CTAGACACTCTA-AAAAAGGCGTGGTGCAGCGGCTGGTGGG 285
||||| 49420 AGTCTCAGGACTTCTAGAGAAAGGCGTGGTGCAGGAGCTTCTGGG 49373

RESULT 9
AX336730
LOCUS AX336730 454 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 7239 from Patent WO0194629.
ACCESSION AX336730
VERSION AX336730.1 GI:18127449
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrihan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 7239 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1..454
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 128 a 113 c 116 g 97 t
ORIGIN

Query Match 52.7%; Score 158; DB 6; Length 454;
Best Local Similarity 98.9%; Pred. No. 1.5e-31;
Matches 180; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATCATGCTTCAGACAACTCCGAGGAGGAGCAAAATCCGGACCTTGTCAAGGATATGT 60
||||| 274 ATCATGCTTCAGACAACTCCGAA-GCAGAGAAATCCGGACCTTGTCAAGGATATGT 332
||||| 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTGTGACAGCAGGAGG 120
||||| 333 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTGTGACAGCAGGAGG 392
||||| 121 CACATGCCAAGCTGGATACCTTGTGATGGAGATCAACACCGCGGACTTTCCTCA 180
||||| 393 CACATG-CAAGCTGGATACCTTGTGATGGAGATCAACACCGCGGACTTTCCTCA 451
||||| 181 CA 182
||||| 452 CA 453

RESULT 10
AB097169
LOCUS AB097169 933 bp DNA linear VRT 09-MAY-2003
DEFINITION Xenopus laevis Psf2 gene, complete cds.
ACCESSION AB097169
VERSION AB097169.1 GI:29365480
KEYWORDS Xenopus laevis (African clawed frog)
SOURCE Xenopus laevis
ORGANISM Xenopus laevis
REFERENCE
AUTHORS Kubota,Y., Takase,Y., Komori,Y., Hashimoto,Y., Arata,T.,

```

# TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL

Kamimura,Y., Araki,H. and Takisawa,H.  
A novel ring-like complex of Xenopus proteins essential for the  
initiation of DNA replication  
Genes Dev. 17 (9), 1141-1152 (2003)  
22615695  
12730133  
2 (bases 1 to 933)  
Kubota,Y. and Takisawa,H.  
Direct Submission  
Submitted (28-NOV-2002) Yumiko Kubota, Osaka University, Department  
of Biology, Graduate School of Science, Machikaneyamacho 1-1,  
Toyonaka, Osaka 560-0043, Japan  
(E-mail:ykubota@bio.sci.osaka-u.ac.jp, Tel:81-6-6850-5554,  
Fax:81-6-6850-5554)  
Location/Qualifiers  
1..933  
/organism="Xenopus laevis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8355"  
21..578  
/gene="Psf2"  
21..578  
/gene="Psf2"  
/function="essential for DNA replication"  
/codon\_start=1  
/product="Psf2"  
/protein\_id="BAC66459.1"  
/db\_xref="GI:29365481"  
/translation="MDASEVEFLAEKEQVTVIPNSLDKVIYLGDLGPNFSLPVEV  
PLMLAINLQROKCRIVPPWMDVEKLEAIRDOERREETFTPMSPYYMLKLLLNH  
AADNIPKADEIETLVKDWTDTRIAKLRSADSFVRQEAHAKLDNLTLMEINTIGTFF  
TESLNHYKLRTSLQNPERSQSDY"  
BASE COUNT 282 a 200 c 208 g 243 t  
ORIGIN

## FEATURES source

gene  
CDS

Query Match 45.5%; Score 136.4; DB 5; Length 933;  
Best Local Similarity 72.7%; Pred. No. 8.1e-26;  
Matches 176; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1 ATCATGCTTCAGACAACTCCGAGGAGGAGCAAAATCCGGACCTTGTCAAGGATATGT 60  
||||| 322 ATCATGCTTCAGACAACTATCCCTTAAGCAGATGAGATTCTGACACTTGTAAAGACACCT 381  
QY 61 GGGACACTCGTATAGCCAAACTCCGAGTGTCTGTGACAGCTTGTGACAGCAGCAGGAG 120  
||||| 382 GGGATACAAGATAGCAAAACTCGGCTCTCTGCTGACAGCTTGTGAAGGGCAGGAAG 441  
QY 121 CACATGCCAAGCTGGATACCTTGTGATGGAGATCAACACCGCGGACTTTCCTCA 180  
||||| 442 CTGATGCCAAGCTGGATAACCTGACGCTAATGGAATTAACACCATCGGAACATTTT 501  
QY 181 CACAAGCGCTCAACACATGTACAAACTCCGACCAACCTCCAGCCTCTGGAAGACCTC 240  
||||| 502 CTGAGTCTTTAAACCATGATACAGCTACCTAGTACCTGACAGAGACCCAGAGGAGGAC 561  
QY 241 AG 242  
||||| 562 AG 563

RESULT 11  
AC009068/c  
LOCUS AC009068 160356 bp DNA linear PRI 27-APR-2001  
DEFINITION Homo sapiens chromosome 16 clone RP11-314K3, complete sequence.  
ACCESSION AC009068  
VERSION AC009068.10 GI:13811892  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 160356)  
DOE Joint Genome Institute and Stanford Human Genome Center.





FEATURES	SOURCE	Location/Qualifiers
55707		55806: gap of unknown length
55807		60092: contig of 4286 bp in length
60093		60192: gap of unknown length
60193		63348: contig of 3156 bp in length
63349		63448: gap of unknown length
63449		69966: contig of 3518 bp in length
69967		67066: gap of unknown length
67067		71746: contig of 4680 bp in length
71747		71846: gap of unknown length
71847		74430: contig of 2584 bp in length
74431		74530: gap of unknown length
74531		79515: contig of 4985 bp in length
79516		79615: gap of unknown length
79616		83377: contig of 3762 bp in length
83378		83477: gap of unknown length
83478		88208: contig of 4731 bp in length
88209		88308: gap of unknown length
88309		93320: contig of 5012 bp in length
93321		93420: gap of unknown length
93421		97345: contig of 3925 bp in length
97346		97445: gap of unknown length
97446		100893: contig of 3448 bp in length
100894		100993: gap of unknown length
100994		107097: contig of 6104 bp in length
107098		107197: gap of unknown length
111968		111963: contig of 4766 bp in length
111969		112063: gap of unknown length
112064		116267: contig of 4204 bp in length
116268		116367: gap of unknown length
116368		123219: contig of 6852 bp in length
123220		123319: gap of unknown length
123320		245210: contig of 121891 bp in length.
		Location/Qualifiers
		1..245210
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="16"
		/clones="CTD-2139A24"
		/clone_lib="Caltech human BAC library D"
BASE COUNT	62851 a 55678 c 56421 g 65631 t 4629 others	
ORIGIN		
Query Match	42.7%;	Score 128.2; DB 2; Length 245210;
Best Local Similarity	97.7%;	Pred. No. 6.1e-24;
Matches 130;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	5	TGCTTCAGACAACTCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGTGGGA 64
Db	115599	TGCTTCAGACAACTCCGGAAGGCAGACGAAATCCGGACCTGGTCAAGGATATGTGGGA 115658
Qy	65	CACTCGTATAGCCAAATCCGAGTGCTGTCGACAGCTTTCTGACACAGCAGGAGGCACA 124
Db	115659	CACTCGTATAGCCAAATCCGAGTGCTGTCGACAGCTTTCTGACACAGCAGGAGGCACA 115718
Qy	125	TGCCAAGCTGGAT 137
Db	115719	TGCCAAGCTAGGT 115731
RESULT 14		
AK091519		
LOCUS	AK091519	3097 bp mRNA linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ34200 fis, clone FCBBF3019663.	
ACCESSION	AK091519	
VERSION	AK091519.1	GI:21749909
KEYWORDS	oligo capping; fis (full insert sequence).	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1	
AUTHORS	Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.	



- \* is not known and their order in this sequence record is
- \* arbitrary. Gaps between the contigs are represented as
- \* runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence.
- \* as soon as it is available and the accession number will
- \* be preserved.

1	2685:	contig of 2685 bp in length
*	2686	2785: gap of 100 bp
*	2786	3775: contig of 990 bp in length
*	3776	3875: gap of 100 bp
*	3876	112935: contig of 103060 bp in length
*	112936	113035: gap of 100 bp
*	113036	115741: contig of 2706 bp in length
*	115742	115841: gap of 100 bp
*	115842	130109: contig of 14368 bp in length
*	130110	130209: gap of 100 bp
*	130210	138730: contig of 8521 bp in length
*	138731	138830: gap of 100 bp
*	138831	186703: contig of 47873 bp in length
*	186704	186803: gap of 100 bp
*	186804	188292: contig of 1489 bp in length.

## FEATURES

```

1. .188292
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /clone="RP24-414A22"
   /clone_lib="RPCI-22 Male Mouse BAC"

```

BASE COUNT	43856 a	49637 c	49457 g	44640 t	702 others
ORIGIN					

	Query Match	33.2%	Score 99.6	DB 2	Length 188292
	Best Local Similarity	72.5%	Pred. No. 2.9e-16		
	Matches 129	Conservative 0	Mismatches 49	Indels 0	Gaps 0
QY	5	TGCTTCAGACAACTCCCGAAGCGAGAGAAATCCGAGCCCTGGTCAAGGATATCTGGGA	64		
Db	55088	TGCTTCTGACACATCCCAAGCAGACACACCATCCGGACACTGATCAAGACCTGTGGGA	55029		
QY	65	CACCTCGTATAGCCAAACTCCGAGTGTCTGCTGACAGCTTTGTGACACAGCAGGAGGCACA	124		
Db	55028	CACACGCATGGCCAAAGCTTCGGAGTGTCTGCTGACAGTTTGTGCGGCAGCAGAGGCACA	54969		
QY	125	TGCCAAGCTGGATAACTTGCACCTTGTATGGAGATCAACACCGAGCGGACTTTCCCTCACA	182		
Db	54968	TGCCAAGGTAGGCATGGTTTCTCTGTCCCTGCCGGTCCGACTGAGTGGGCACTGCCTGCCA	54911		

Search completed: October 14, 2003, 17:02:31  
Job time : 1479 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 11:32:16 ; Search time 1272.67 Seconds  
(without alignments)  
5729.185 Million cell updates/sec

Title: US-09-854-124-6  
Perfect score: 300  
Sequence: 1 aattccytgtctgtgga.....cgaaaggaagaactttac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286.4	95.5	959	10	BG118214
2	285	95.0	714	10	BG104289
3	284.8	94.9	388	9	AI914017
4	284.8	94.9	429	10	BE262182
					BG118214 602350003
					BG104289 602310878
					AI914017 wd74803.x
					BE262182 601150452

5	284.8	94.9	487	9	AI827298
6	284.8	94.9	488	9	AA251319
7	284.8	94.9	515	10	BF312015
8	284.8	94.9	532	9	AI583174
9	284.8	94.9	550	9	AI828992
10	284.8	94.9	573	10	BE514978
11	284.8	94.9	671	10	BE258240
12	284.8	94.9	696	9	AW249012
13	284.8	94.9	702	9	AV712739
14	284.8	94.9	716	12	BG825252
15	284.8	94.9	729	10	BE251621
16	284.8	94.9	731	10	BG475509
17	284.8	94.9	736	10	BE312319
18	284.8	94.9	740	13	BU595469
19	284.8	94.9	749	12	BG772776
20	284.8	94.9	761	10	BG339264
21	284.8	94.9	761	12	BI257993
22	284.8	94.9	771	10	BG475384
23	284.8	94.9	833	13	BU597296
24	284.8	94.9	834	10	BG527542
25	284.8	94.9	877	13	BU940719
26	284.8	94.9	882	10	BF684568
27	284.8	94.9	891	11	BC022839
28	284.8	94.9	897	10	BG519560
29	284.8	94.9	905	9	AI525760
30	284.8	94.9	909	13	BQ229290
31	284.8	94.9	914	14	CA455226
32	284.8	94.9	943	10	BF683514
33	284.8	94.9	982	13	BU957713
34	284.8	94.9	997	10	BG420251
35	284.8	94.9	1007	10	BE795838
36	284.8	94.9	1030	10	BF311745
37	284.8	94.9	1057	10	BF312439
38	284.8	94.9	1067	13	BX415204
39	284.8	94.9	1182	11	AF125098
40	284.8	94.9	1193	9	AL529785
41	284	94.7	624	10	BG284180
42	284	94.7	908	10	BE267221
43	284	94.7	1201	9	AL560880
44	283.2	94.4	568	10	AW958853
45	283.2	94.4	683	10	BE260083

#### ALIGNMENTS

RESULT 1  
BG118214

LOCUS  
DEFINITION

BG118214 602350003f1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:444801 5', mRNA linear EST 30-JAN-2001

ACCESSION  
BG118214

VERSION  
BG118214.1 GI:12611720

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 959)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cga@r@mail.nih.gov](mailto:cga@r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAMI0221 row: f column: 02

High quality sequence stop: 719.





/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2422723"  
 /tissue\_type="squamous cell carcinoma, poorly  
 differentiated (4 pooled tumors, including primary and  
 metastatic)"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Lu19"  
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; 1st strand cDNA was prepared from a  
 pooled lung tumor tissue, and was then primed with a Not I  
 - oligo(dT) primer. Double-stranded cDNA was ligated to  
 Eco RI adaptors (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of the modified  
 pT73 vector. Library went through one round of  
 normalization. Library constructed by Bento Soares and M.  
 Fatima Bonaldo."

BASE COUNT 119 a 132 c 138 g 98 t

Query Match 94.9%; Score 284.8; DB 9; Length 487;

Best Local Similarity 99.3%; Pred. No. 1.8e-66;

Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGGAGGCTGTGGCTGCGAGCTGCGCGCCGCGCATGGAGCTGCGAGGTGCAATTCCT 72

Db 43 GACCGTGGAGGCTGTGGCTGCGAGCTGCGCGCCGCGCATGGAGCTGCGAGGTGCAATTCCT 102

QY 73 CGCGGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAAGATCTACCTCAT 132

Db 103 CGCGGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAAGATCTACCTCAT 162

QY 133 CGGGGGGACCTGGGGCCCTTTAAACCTGCTTTACCCGCGAGTGGCCCTGTGGCTGCG 192

Db 163 CGGGGGGACCTGGGGCCCTTTAAACCTGCTTTACCCGCGAGTGGCCCTGTGGCTGCG 222

QY 193 GATTAACCTGAAACAAGACAGAAATGTCCTGCTCCCTCCAGAGTGGATGGATGAGA 252

Db 223 GATTAACCTGAAACAAGACAGAAATGTCCTGCTCCCTCCAGAGTGGATGGATGAGA 282

QY 253 AAGTTGGAGAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 300

Db 283 AAGTTGGAGAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 330

RESULT 6

AA251319

LOCUS zs04d09.s1 NCI\_CGAP\_GCB1 488 bp mRNA linear EST 13-AUG-1997

DEFINITION similar to SW:YJH2\_YEAST P40359 HYPOTHETICAL 25.1 KD PROTEIN IN

SCPI60-MRPL8 INTERGENIC REGION. ;, mRNA sequence.

ACCESSION AA251319

VERSION AA251319

KEYWORDS EST.

SOURCE AA251319.1 GI:1886282

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 488)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

Insert length: 910 Std Error: 0.00

Seq primer: -41ml3 fwd. Et from Amersham

# FEATURES

## source

High quality sequence stop: 463.

Location/Qualifiers

1. 488

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:684209"

/tissue\_type="germinal center B cell"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP\_GCB1"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD20+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 [5'-GTTACCAATCTGAAGTGGGCGCCGCTCATTTTTTTTTTTT-3'  
 ]. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 118 a 134 c 138 g 98 t

Query Match

Best Local Similarity 94.9%; Score 284.8; DB 9; Length 488;

Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGGAGGCTGTGGCTGCGAGCTGCGCGCCGCGCATGGAGGTGCAATTCCT 72

Db 43 GACCGTGGAGGCTGTGGCTGCGAGCTGCGCGCCGCGCATGGAGGTGCAATTCCT 102

QY 73 CGCGGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAAGATCTACCTCAT 132

Db 103 CGCGGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAAGATCTACCTCAT 162

QY 133 CGGGGGGACCTGGGGCCCTTTAAACCTGCTTTACCCGCGAGTGGCCCTGTGGCTGCG 192

Db 163 CGGGGGGACCTGGGGCCCTTTAAACCTGCTTTACCCGCGAGTGGCCCTGTGGCTGCG 222

QY 193 GATTAACCTGAAACAAGACAGAAATGTCCTGCTCCCTCCAGAGTGGATGGATGAGA 252

Db 223 GATTAACCTGAAACAAGACAGAAATGTCCTGCTCCCTCCAGAGTGGATGGATGAGA 282

QY 253 AAGTTGGAGAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 300

Db 283 AAGTTGGAGAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 330

RESULT 7

BF312015

LOCUS 601897874F1 NIH\_MGC\_19 515 bp mRNA linear EST 21-NOV-2000

DEFINITION mRNA sequence.

ACCESSION BF312015

VERSION BF312015

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 515)

NIH-MGC http://mgc.ncbi.nlm.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCMI018 row: c column: 19  
High quality sequence stop: 515.  
Location/Qualifiers  
1. 515

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4127178"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_19"  
/note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 132 a 134 c 146 g 103 t  
ORIGIN  
Query Match 94.9%; Score 284.8; DB 10; Length 515;  
Best Local Similarity 99.3%; Pred. No. 1.9e-66;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 GTCGTGAGGCTCTGGCTGACGCTCGCGCCGCATGGACGCTGCGAGGTGCAATTCTCT 72  
DB 26 GACCGTAGGCTCTGGCTGACGCTCGCGCCGCATGGACGCTGCGAGGTGCAATTCTCT 85  
QY 73 CGCCGAGAAGGAGCTGGTTACCATTCATCCCAACTTCAGTCTGGACAAAGATCTACTCAT 132  
DB 86 CGCCGAGAAGGAGCTGGTTACCATTCATCCCAACTTCAGTCTGGACAAAGATCTACTCAT 145  
QY 133 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTGGAGTGGCCCTGTGGCTGGC 192  
DB 146 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTGGAGTGGCCCTGTGGCTGGC 205  
QY 193 GATTAACCTGAAACAAAGACAGAAATGTCGCCTGCTCCCTCCAGAGTGGATGTAGTAGA 252  
DB 206 GATTAACCTGAAACAAAGACAGAAATGTCGCCTGCTCCCTCCAGAGTGGATGTAGTAGA 265  
QY 253 AAAGTTGGAGAAGATGAGGAGTATCAATCAACGAAGGAAGAACTTTTAC 300  
DB 266 AAAGTTGGAGAAGATGAGGAGTATCAATCAACGAAGGAAGAACTTTTAC 313

RESULT 8  
AI583174  
LOCUS  
DEFINITION  
tq64e05.x1 NCI-CGAP\_Lu19 Homo sapiens cDNA clone IMAGE:2213600 3', similar to SW:YJH2\_YEAST P40359 HYPOPHYSICAL 25.1 KD PROTEIN IN SMC3-MRPL8 INTERGENIC REGION. [1] ;, mRNA sequence.  
AI583174  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 532)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.

AI583174.1 GI:4569071  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 532)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/dbbrp/image/image.html](http://www-bio.llnl.gov/dbbrp/image/image.html)  
Insert Length: 1092 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 459  
POLYA=No.

FEATURES  
source

Location/Qualifiers  
1. 532  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2213600"  
/tissue\_type="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"  
/dev\_stage="adult"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NCI-CGAP\_Lu19"  
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 140 c 149 g 104 t 1 others  
ORIGIN  
Query Match 94.9%; Score 284.8; DB 9; Length 532;  
Best Local Similarity 99.3%; Pred. No. 1.9e-66;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 GTCGTGAGGCTCTGGCTGACGCTCGCGCCGCATGGACGCTGCGAGGTGCAATTCTCT 72  
DB 25 GACCGTAGGCTCTGGCTGACGCTCGCGCCGCATGGACGCTGCGAGGTGCAATTCTCT 84  
QY 73 CGCCGAGAAGGAGCTGGTTACCATTCATCCCAACTTCAGTCTGGACAAAGATCTACTCAT 132  
DB 85 CGCCGAGAAGGAGCTGGTTACCATTCATCCCAACTTCAGTCTGGACAAAGATCTACTCAT 144  
QY 133 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTGGAGTGGCCCTGTGGCTGGC 192  
DB 145 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTGGAGTGGCCCTGTGGCTGGC 204  
QY 193 GATTAACCTGAAACAAAGACAGAAATGTCGCCTGCTCCCTCCAGAGTGGATGTAGTAGA 252  
DB 205 GATTAACCTGAAACAAAGACAGAAATGTCGCCTGCTCCCTCCAGAGTGGATGTAGTAGA 264  
QY 253 AAAGTTGGAGAAGATGAGGAGTATCAATCAACGAAGGAAGAACTTTTAC 300  
DB 265 AAAGTTGGAGAAGATGAGGAGTATCAATCAACGAAGGAAGAACTTTTAC 312

RESULT 9  
AI828992  
LOCUS  
DEFINITION  
wj48f05.x1 NCI-CGAP\_Lu19 Homo sapiens cDNA clone IMAGE:2406081 3', similar to WP:F31C3.5 CE17734 ;, mRNA sequence.  
AI828992  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 550)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index



REFERENCE 1 (bases 1 to 671)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/dbp/image/image.html](http://www.bio.llnl.gov/dbp/image/image.html)  
 Plate: LLC161 row: i column: 24  
 High quality sequence stop: 655.  
 Location/Qualifiers  
 FEATURES  
 source  
 1..671  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone IMAGE:3356255  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_16"  
 /note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
 BASE COUNT 172 a 179 c 184 g 136 t  
 ORIGIN  
 Query Match 94.9%; Score 284.8; DB 10; Length 671;  
 Best Local Similarity 99.3%; Pred. No. 2e-66;  
 Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 13 GTCGGTGAGGCTCTGGCTGCAGCTCGCGCCCGCCATGACGCTGCCGAGGTGCAATTCCT 72  
 Db 8 GACCTTGAGGCTCTGGCTGCAGCTCGCGCCCGCCATGACGCTGCCGAGGTGCAATTCCT 67  
 QY 73 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 132  
 Db 68 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 127  
 QY 133 CGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGGAAGTGCCTGTGGCTGGC 192  
 Db 128 CGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGGAAGTGCCTGTGGCTGGC 187  
 QY 193 GATTAACTGAAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATGAGA 252  
 Db 188 GATTAACTGAAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATGAGA 247  
 QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 300  
 Db 248 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 295  
 RESULT 12  
 AW249012  
 LOCUS AW249012  
 DEFINITION 2821094.5prime NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:2821094 5', mRNA sequence.  
 ACCESSION AW249012  
 VERSION AW249012.1 GI:6592005  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 696)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Other\_Ests: 2821094.3prime  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing project clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/dbp/image/image.html](http://www.bio.llnl.gov/dbp/image/image.html)  
 Scores: PHRED from University of Washington Genome Center  
 Trimming: cross\_match from University of Washington Genome Center  
 PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center  
 http://www.genome.washington.edu  
 Plate: LLCM5 row: O column: 15  
 High quality sequence stop: 573.  
 Location/Qualifiers  
 FEATURES  
 source  
 1..696  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone IMAGE:2821094  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_7"  
 /note="Organ: lung; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 179 a 183 c 192 g 142 t  
 ORIGIN  
 Query Match 94.9%; Score 284.8; DB 9; Length 696;  
 Best Local Similarity 99.3%; Pred. No. 2e-66;  
 Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 13 GTCGGTGAGGCTCTGGCTGCAGCTCGCGCCCGCCATGACGCTGCCGAGGTGCAATTCCT 72  
 Db 13 GACCTTGAGGCTCTGGCTGCAGCTCGCGCCCGCCATGACGCTGCCGAGGTGCAATTCCT 72  
 QY 73 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 132  
 Db 73 CGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 132  
 QY 133 CGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGGAAGTGCCTGTGGCTGGC 192  
 Db 133 CGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGGAAGTGCCTGTGGCTGGC 192  
 QY 193 GATTAACTGAAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATGAGA 252  
 Db 193 GATTAACTGAAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATGAGA 252  
 QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 300  
 Db 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAACTTTTAC 300  
 RESULT 13  
 AV712739  
 LOCUS AV712739  
 DEFINITION AV712739 DCA Homo sapiens cDNA clone DCAACE07 5', mRNA sequence.  
 ACCESSION AV712739  
 VERSION AV712739.1 GI:10732045  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

# REFERENCE AUTHORS

1 (bases 1 to 702)  
Xu, X., Gu, J., Liu, F., Qu, J., Zhao, M., Li, Y., Huang, Q., Zhou, J., Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X., Cheng, Z., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Cheng, Z., and Han, Z.

# TITLE JOURNAL COMMENT

Homo sapiens cDNA clones  
Unpublished  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chc.sh.cn  
This clone is available at CHGC in Shanghai.

# FEATURES source

Location/Qualifiers  
1..702  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DCAACE07"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCA"  
/note="Vector: pTriplex2; Site\_1: sfIIA; Site\_2: sfIIB"  
BASE COUNT 183 a 181 c 184 g 147 t 7 others  
ORIGIN

Query Match 94.9%; Score 284.8; DB 9; Length 702;  
Best Local Similarity 99.3%; Pred. No. 2e-66;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGGAGCTGCGCTGAGCTCGCGCGCCATGCGAGTGGCCGAGTGGCAATTCCT 72  
DB 5 GACCGTGGAGCTGCGCTGAGCTCGCGCGCCATGCGAGTGGCCGAGTGGCAATTCCT 64  
QY 73 CGCGGAGAAGGAGCTGTTACCATTTATCCCAACTTCAGTCTGACAAAGATCTACCTCAT 132  
DB 65 CGCGGAGAAGGAGCTGTTACCATTTATCCCAACTTCAGTCTGACAAAGATCTACCTCAT 124  
QY 133 CGGGGGGAGCTGGGGCTTTTAACTTGTTCACCGTGGAGTGGCCCTGTGGCGGCG 192  
DB 125 CGGGGGGAGCTGGGGCTTTTAACTTGTTCACCGTGGAGTGGCCCTGTGGCGGCG 184  
QY 193 GATTAACTTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGTAGTA 252  
DB 185 GATTAACTTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGTAGTA 244  
QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAAGAACTTTTAC 300  
DB 245 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAAGAACTTTTAC 292

RESULT 14  
BG825252  
LOCUS 716 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602747675F1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:4900602 5',  
mRNA sequence.

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

BG825252  
BG825252.1 GI:14172839  
EST.  
Homo sapiens (human)  
Homo sapiens

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 716)  
NTH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLCMI794 row: e column: 19

High quality sequence stop: 716.

Location/Qualifiers

1..716

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4900602"

/tissue\_type="rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_17"

/note="Organ: muscle; Vector: pOTB7; Site\_1: EcoRI;

Site\_2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCAGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 186 a 185 c 198 g 147 t

ORIGIN

# Query Match

Best Local Similarity 94.9%; Score 284.8; DB 12; Length 716;

Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGGAGCTGCGCTGAGCTCGCGCGCCATGCGAGTGGCCGAGTGGCAATTCCT 72  
DB 15 GACCGTGGAGCTGCGCTGAGCTCGCGCGCCATGCGAGTGGCCGAGTGGCAATTCCT 74  
QY 73 CGCGGAGAAGGAGCTGTTACCATTTATCCCAACTTCAGTCTGACAAAGATCTACCTCAT 132  
DB 75 CGCGGAGAAGGAGCTGTTACCATTTATCCCAACTTCAGTCTGACAAAGATCTACCTCAT 134  
QY 133 CGGGGGGAGCTGGGGCTTTTAACTTGTTCACCGTGGAGTGGCCCTGTGGCGGCG 192  
DB 135 CGGGGGGAGCTGGGGCTTTTAACTTGTTCACCGTGGAGTGGCCCTGTGGCGGCG 194  
QY 193 GATTAACTTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGTAGTA 252  
DB 195 GATTAACTTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGTAGTA 254  
QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAAGAACTTTTAC 300  
DB 255 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAAGAACTTTTAC 302

# RESULT 15 BE251621 LOCUS

DEFINITION 601110640F1 NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:3351416 5',  
mRNA sequence.

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

BE251621  
BE251621.1 GI:9121737  
EST.  
Homo sapiens (human)  
Homo sapiens

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 729)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLC148 row: p column: 09  
High quality sequence stop: 628.  
Location/Qualifiers  
1. .729  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3351416"  
/tissue\_type="retinoblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_16"  
/note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 181 a 193 c 214 g 141 t  
ORIGIN  
Query Match 94.9%; Score 284.8; DB 10; Length 729;  
Best Local Similarity 99.3%; Pred. No. 2e-66;  
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 GTCGGTGAGGCTGTGGCCCTGCAGCTCGCGCCGACATGACGCTGCCGAGGTGGAATTCCT 72  
Db 8 GACCGTGGGCTCTGSCCTGCAGCTCGCGCCGACATGACGCTGCCGAGGTGGAATTCCT 67  
QY 73 CGCCGAGAGGAGCTGGTTACCATTTATCCCACTTCAGTCTGGACAAGATCTACCTCAT 132  
Db 68 CGCCGAGAGGAGCTGGTTACCATTTATCCCACTTCAGTCTGGACAAGATCTACCTCAT 127  
QY 133 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTTGGAAGTCCCTGTGGCTGGC 192  
Db 128 CGGGGGGACCTGGGGCCCTTTAAACCTGGTTTACCCGTTGGAAGTCCCTGTGGCTGGC 187  
QY 193 GATTAACCTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGGATGTAGA 252  
Db 188 GATTAACCTGAACAAAGACAGAAATGTCGCTGCTCCCTCCAGAGTGGATGGATGTAGA 247  
QY 253 AAAGTTGGAGAGATGAGGGATCATGAACGAAGGAAGAACTTTTAC 300  
Db 248 AAAGTTGGAGAGATGAGGGATCATGAACGAAGGAAGAACTTTTAC 295

Search completed: October 14, 2003, 14:09:09  
Job time : 1278.67 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 11:32:16 ; Search time 1272.67 Seconds  
(without alignments)  
5729.185 Million cell updates/sec

Title: US-09-854-124-5

Perfect score: 300

Sequence:

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

al number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estlin:\*\*

4: em\_estnu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_hic:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_hic:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pin:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rod:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vri:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	300	100.0	533	14	CB112523 K-EST0154
C 2	300	100.0	559	12	BM126492 if06h06.y
C 3	300	100.0	659	13	BU729618 UT-E-CQ1-
C 4	300	100.0	716	10	BE382866 601297755

5	300	100.0	801	12	BI196248
6	300	100.0	876	10	BI181130
7	300	100.0	896	10	BE562088
8	300	100.0	914	10	BE547290
9	300	100.0	943	10	BF683514
10	300	100.0	970	9	AL560669
11	300	100.0	1007	12	BM449472
12	300	100.0	1064	13	BQ277667
13	300	100.0	1071	13	BU184963
14	300	100.0	1071	9	AL560880
15	300	100.0	1201	9	AL582250
16	299.6	99.9	1201	9	AL582217
17	299	99.7	1810	12	BQ009838
18	298.8	99.6	1193	9	AL529785
19	298.4	99.5	488	14	CB160336
20	298.4	99.5	696	9	AW249012
21	298.4	99.5	712	13	BU625683
22	298.4	99.5	714	10	BG104289
23	298.4	99.5	716	12	BG825252
24	298.4	99.5	730	10	BE296429
25	298.4	99.5	790	10	BE795306
26	298.4	99.5	866	10	BF795157
27	298.4	99.5	867	13	BQ233393
28	298.4	99.5	891	11	BC022839
29	298.4	99.5	921	10	BE799212
30	298.4	99.5	932	10	BE796384
31	298.4	99.5	1028	10	BE561044
32	298.4	99.5	1030	10	BF311745
33	298.4	99.5	1182	11	AF125098
34	298	99.3	833	13	BU597296
35	297.4	99.1	494	14	H64860
36	296.8	98.9	895	13	BU539659
37	293.2	97.7	1056	13	BX456910
38	292.4	97.5	667	9	AU126087
39	292	97.3	966	9	AL582077
40	291.4	97.1	1201	9	AL562756
41	290	96.7	649	10	BE514071
42	289.6	96.5	1007	10	BE795838
43	288.6	96.2	1201	9	AL526847
44	288	96.0	761	12	BI257993
45	287.4	95.8	966	10	BE791539

ALIGNMENTS

RESULT 1  
CB112523/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB112523  
K-EST0154352 L6ChoCK0 Homo sapiens CDNA clone L6ChoCK0-4-D04 5',  
533 bp mRNA linear EST 28-JAN-2003  
mRNA sequence.

CB112523  
GI:27938330

EST.  
Homo sapiens (human)

CB112523  
Homo sapiens

CB112523  
Eukaryota: Metazoa; Chordata; Vertebrata; Euteleostomi;

CB112523  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CB112523  
1 (bases 1 to 533)

CB112523  
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

CB112523  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

CB112523  
Kim, Y.S.

CB112523  
21C Frontier Korean EST Project 2001

CB112523  
Unpublished

CB112523  
Contact: Kim YS

CB112523  
Genome Research Center

CB112523  
Korea Research Institute of Bioscience & Biotechnology

CB112523  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

CB112523  
Tel: +82-42-860-4470

CB112523  
Fax: +82-42-860-4409

CB112523  
Email: yongsung@mail.kribb.re.kr

CB112523  
Plate: 4 row: D column: 04

CB112523  
High quality sequence stop: 533.





LOCUS BU729618 659 bp mRNA linear EST 09-OCT-2002  
 DEFINITION UI-E-CQ1-ay-b-18-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone  
 UI-E-CQ1-ay-b-18-0-UI 3', mRNA sequence.  
 ACCESSION BU729618  
 VERSION BU729618.1 GI:23652681  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 659)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 FORWARD  
 POLYA-yes.

## FEATURES

source Location/Qualifiers  
 1..659  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-CQ1-ay-b-18-0-UI"  
 /tissue\_type="optic nerve"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-CQ1"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-E-CQ1 is a normalized cDNA library containing the  
 following tissue(s): optic nerve. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dr primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 CCATTAGTG. This library was created for the program, Eye  
 discovery in the Visual System, supported by National Eye  
 Institute (NEI).  
 TAG LIB-UI-E-CQ1  
 TAG\_TISSUE=human optic nerve  
 TAG\_SEQ=CCATTAGTG"  
 149 a 151 g 200 t

## BASE COUNT

ORIGIN  
 Query Match 100.0%; Score 300; DB 13; Length 659;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 |||||||  
 DB 509 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 450  
 |||||||  
 QY 61 CTGCTGACAGCTTTGTGAGACACAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 120

DB 449 CTGCTGACAGCTTTGTGAGACACAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 390  
 |||||||  
 QY 121 TGGAGATCAACACACAGCGGACTTCTCACACAGCGCTCAACACACATGTACAACTCC 180  
 |||||||  
 DB 389 TGGAGATCAACACACAGCGGACTTCTCACACAGCGCTCAACACACATGTACAACTCC 330  
 |||||||  
 QY 181 GCACGAACCTCCAGCCCTCTGGAGAGTACTCACTAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 |||||||  
 DB 329 GCACGAACCTCCAGCCCTCTGGAGAGTACTCACTAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 270  
 |||||||  
 QY 241 GCAGCGGCTTCTGGGGGATGTAGCGCTCAGGACGTAGGAGTACTCTGGTTCTGG 300  
 |||||||  
 DB 269 GCAGCGGCTTCTGGGGGATGTAGCGCTCAGGACGTAGGAGTACTCTGGTTCTGG 210  
 |||||||

## RESULT 4

BE382866 716 bp mRNA linear EST 21-JUL-2000  
 LOCUS 601297755F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3627852 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE382866  
 VERSION BE382866.1 GI:9328231  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 716)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC

## FEATURES

source Location/Qualifiers  
 1..716  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3627852"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; CDNA made by oligo-dr priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5',  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 185 a 179 c 195 g 157 t

## BASE COUNT

ORIGIN  
 Query Match 100.0%; Score 300; DB 10; Length 716;  
 Best Local Similarity 100.0%; Pred. No. 2e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 |||||||  
 DB 337 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 396  
 |||||||  
 QY 61 CTGCTGACAGCTTTGTGAGACACAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 120  
 |||||||  
 DB 397 CTGCTGACAGCTTTGTGAGACACAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 456  
 |||||||

121 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 180  
 457 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 516  
 181 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCTGGT 240  
 517 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCTGGT 576  
 241 GCAGCGGCTTCTCTGGGGATGTAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTTGG 300  
 577 GCAGCGGCTTCTCTGGGGATGTAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTTGG 636

RESULT 5  
 BI196248  
 LOCUS  
 DEFINITION 602754709F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4890147 5',  
 mRNA sequence.  
 ACCESSION BI196248  
 VERSION  
 WORDS  
 EST. BI196248.1 GI:14651268  
 ORGANCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS 1 (bases 1 to 801)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: L1CMI767 row: b column: 04  
 High quality sequence stop: 781.

FEATURES  
 source  
 location/Qualifiers  
 1..801  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4890147"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
 Note: this is a NIH\_MGC Library."

BASE COUNT 208 a 196 c 216 g 181 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 12; Length 801;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 DB 356 ACAGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 415  
 QY 61 CTGCTCACAGCTTTGTGACACAGCAGGACATGCCAAGCTGGATATACTTGACCTTGA 120  
 DB 416 CTGCTCACAGCTTTGTGACACAGCAGGAGGACATGCCAAGCTGGATATACTTGACCTTGA 475  
 QY 121 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 180

Db 476 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 535  
 QY 181 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCTGGT 240  
 Db 536 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCTGGT 595  
 QY 241 GCAGCGGCTTCTCTGGGGATGTAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTTGG 300  
 Db 596 GCAGCGGCTTCTCTGGGGATGTAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTTGG 655

RESULT 6  
 BG181130  
 LOCUS  
 DEFINITION 602329264F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4430823 5',  
 mRNA sequence.  
 ACCESSION BG181130  
 VERSION  
 WORDS  
 EST. BG181130.1 GI:12687833  
 ORGANCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS 1 (bases 1 to 876)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM10184 row: o column: 16  
 High quality sequence stop: 674.

FEATURES  
 source  
 location/Qualifiers  
 1..876  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4430823"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies."  
 Note: this is a NIH\_MGC Library."

BASE COUNT 246 a 217 c 214 g 199 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 10; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db 115 ACAGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 174  
 QY 61 CTGCTCACAGCTTTGTGACACAGCAGGACATGCCAAGCTGGATATACTTGACCTTGA 120  
 Db 175 CTGCTCACAGCTTTGTGACACAGCAGGACATGCCAAGCTGGATATACTTGACCTTGA 234  
 QY 121 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 180  
 Db 235 TGGAGATCAACACCGGGGACTTTCTCCACACAGCGCTCAACACATGTACAACCTCC 294  
 QY 181 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGGT 240

Db 295 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 354  
 QY 241 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 Db 355 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 414

RESULT 7  
 BE562088 896 bp mRNA linear EST 15-AUG-2000  
 LOCUS 601345038F1 NIH\_MCC\_8 Homo sapiens cDNA clone IMAGE:3678080 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE562088  
 VERSION BE562088.1 GI:9805808  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LCM354 row: k column: 09  
 High quality sequence start: 24  
 High quality sequence stop: 840.

FEATURES  
 source  
 1..896  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3678080"  
 /tissue\_type="Burkitt lymphoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_8"  
 /note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 218 a 230 c 251 g 197 t

Query Match 100.0%; Score 300; DB 10; Length 896;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCTTGTGACAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db 438 ACGAAATCCGACCTTGTGACAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 497  
 QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACTTGACCTTGA 120  
 Db 498 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACTTGACCTTGA 557  
 QY 121 TGGAGATCAACACACGCGGACTTTCCTCACACAAGCGCTCAACCATGTACAACTCC 180  
 Db 558 TGGAGATCAACACACGCGGACTTTCCTCACACAAGCGCTCAACCATGTACAACTCC 617  
 QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240

Db 618 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 677  
 QY 241 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 Db 678 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 737

RESULT 8  
 BE547290 914 bp mRNA linear EST 09-AUG-2000  
 LOCUS 601073715F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:3460010 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE547290  
 VERSION BE547290.1 GI:9775935  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: AFCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAMB453 row: m column: 03  
 High quality sequence stop: 658.

FEATURES  
 source  
 1..914  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3460010"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.4 kb. Library prepared by Life  
 Technologies."  
 254 a 227 c 238 g 195 t

Query Match 100.0%; Score 300; DB 10; Length 914;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCTTGTGACAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db 104 ACGAAATCCGACCTTGTGACAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 163  
 QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACTTGACCTTGA 120  
 Db 164 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACTTGACCTTGA 223  
 QY 121 TGGAGATCAACACACGCGGACTTTCCTCACACAAGCGCTCAACCATGTACAACTCC 180  
 Db 224 TGGAGATCAACACACGCGGACTTTCCTCACACAAGCGCTCAACCATGTACAACTCC 283  
 QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 Db 284 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 343  
 QY 241 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 Db 344 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 403

*Cervical  
 cell line*

RESULT 9  
BF683514  
LOCUS  
DEFINITION  
602139714F1 NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4300947 5',  
mRNA sequence.  
ACCESSION  
BF683514  
VERSION  
BF683514.1 GI:11968922  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 943)  
AUTHORS  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L10M1160 row: 1 column: 04  
High quality sequence stop: 709.

FEATURES  
source  
1. 943  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4300947"  
/tissue\_type="leiomyosarcoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_46"  
/note="Organ: uterus; Vector: pOTB7; Site:1; XhoI; Site:2;  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."  
BASE COUNT 214 a 253 c 272 g 204 t

Query Match 100.0%; Score 300; DB 10; Length 943;  
Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCAAACTCCGAGTGT 60  
DB 369 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCAAACTCCGAGTGT 428  
QY 61 CTGCTCACAGCTTTGTGAGACAGCAGGACACATGCCAAGCTGGGATTAAGTTCGACCTTGA 120  
DB 429 CTGCTCACAGCTTTGTGAGACAGCAGGACACATGCCAAGCTGGGATTAAGTTCGACCTTGA 488  
QY 121 TGGAGATCAACACACAGCGGAGCTTTCTCACACAAGCGCTCAACACATGTACAACTCC 180  
DB 489 TGGAGATCAACACACAGCGGAGCTTTCTCACACAAGCGCTCAACACATGTACAACTCC 548  
QY 181 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGTGT 240  
DB 549 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGTGT 608  
QY 241 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGTGTAGGATCTCGTGGTTCGTGG 300  
DB 609 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGTGTAGGATCTCGTGGTTCGTGG 668

RESULT 10  
AL560669  
LOCUS  
DEFINITION  
AL560669 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
Homo sapiens cDNA clone CS0DL003YI21 5-PRIME, mRNA sequence.  
ACCESSION  
AL560669  
VERSION  
AL560669.2 GI:31284799  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 970)  
AUTHORS  
Li W.B., Gruber C., Jessee J., and Pollayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
COMMENT  
On Feb 15, 2001 this sequence version replaced gi:12907354.

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7987.r For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL003AE11Q1&cluster=7987.r)  
cgi-bin/cluster.cgi?seq=CS0DL003AE11Q1&cluster=7987.r. Contact :  
Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
Faraday Avenue Genoscope sequence ID : CS0DL003AE11Q1.  
Location/Qualifiers  
1. 970  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DL003YI21"  
/cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
/cell\_line="RAMOS CELL LINE"  
/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 257 a 243 c 251 g 219 t

FEATURES  
source  
1. 970

Query Match 100.0%; Score 300; DB 9; Length 970;  
Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCAAACTCCGAGTGT 60  
DB 419 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCAAACTCCGAGTGT 478  
QY 61 CTGCTCACAGCTTTGTGAGACAGCAGGACACATGCCAAGCTGGGATTAAGTTCGACCTTGA 120  
DB 479 CTGCTCACAGCTTTGTGAGACAGCAGGACACATGCCAAGCTGGGATTAAGTTCGACCTTGA 538  
QY 121 TGGAGATCAACACACAGCGGAGCTTTCTCACACAAGCGCTCAACACATGTACAACTCC 180  
DB 539 TGGAGATCAACACACAGCGGAGCTTTCTCACACAAGCGCTCAACACATGTACAACTCC 598  
QY 181 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGTGT 240  
DB 599 GCACGAACCTCCAGGCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGTGT 658  
QY 241 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGTGTAGGATCTCGTGGTTCGTGG 300  
DB 659 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGTGTAGGATCTCGTGGTTCGTGG 718

RESULT 11  
 BM449472  
 LOCUS  
 DEFINITION AGENCOURT\_6400913 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5493679  
 5', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1007)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LAM12117 row: m column: 08  
 High quality sequence stop: 711.  
 Location/Qualifiers  
 1..1007  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5493679"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_67"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies." 259 a 254 c 262 g 231 t 1 others

BASE COUNT 259 a 254 c 262 g 231 t 1 others  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 12; Length 1007;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCTGTGTCAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 357 ACGAAATCCGACCTGTGTCAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 416  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 120  
 417 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 476  
 QY 121 TGGAGATCAACACACCGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAACTCC 180  
 477 TGGAGATCAACACACCGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAACTCC 536  
 QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 240  
 537 GCACGAACCTCCAGCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 596  
 QY 241 GCAGCGGCTTCTGTTGGGGATGTAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 597 GCAGCGGCTTCTGTTGGGGATGTAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 656

RESULT 12  
 BQ277667  
 LOCUS  
 DEFINITION AGENCOURT\_7048107 NIH\_MGC\_109 Homo sapiens cDNA clone IMAGE:5804514  
 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BQ277667  
 BQ277667.1 GI:20487875  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1064)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LCM2042 row: d column: 19  
 High quality sequence stop: 675.  
 Location/Qualifiers  
 1..1064  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5804514"  
 /tissue\_type="teratocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_109"  
 /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; CDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library." 275 a 268 c 273 g 245 t 3 others

FEATURES  
 source

BASE COUNT 275 a 268 c 273 g 245 t 3 others  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 13; Length 1064;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCTGTGTCAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 364 ACGAAATCCGACCTGTGTCAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 423  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 120  
 424 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 483  
 QY 121 TGGAGATCAACACACCGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAACTCC 180  
 484 TGGAGATCAACACACCGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAACTCC 543  
 QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 240  
 544 GCACGAACCTCCAGCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 603  
 QY 241 GCAGCGGCTTCTGTTGGGGATGTAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 604 GCAGCGGCTTCTGTTGGGGATGTAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 663

RESULT 13  
 BUI84963  
 LOCUS  
 DEFINITION AGENCOURT\_7186943 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6002100  
 5', mRNA sequence.  
 ACCESSION  
 BUI84963

VERSION BUI84963.1 GI:22698947  
 .KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1071)  
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: CGAP (Stanford)  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2269 row: m column: 13  
 High quality sequence stop: 718.  
 Location/Qualifiers  
 1. 1071  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6002100"  
 /tissue\_type="hepatocellular carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_100"  
 /note="Organ: liver; Vector: pORF7; Site: 1; XhoI; Site: 2;  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

URES  
 source

BASE COUNT 283 a 262 c 275 g 246 t 5 others  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 13; Length 1071;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 60  
 Db 359 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 418  
 61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 120  
 Db 419 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 478  
 121 TGGAGATCAACACACAGCGGACTTTCCTCACACAAGCGCTCAACACATGTACAACTCC 180  
 Db 479 TGGAGATCAACACACAGCGGACTTTCCTCACACAAGCGCTCAACACATGTACAACTCC 538  
 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGAGCTCTAGAGAAAGGCTGTGT 240  
 Db 539 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGAGCTCTAGAGAAAGGCTGTGT 598  
 241 GCAGCGGCTTCTCGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTTCTGG 300  
 Db 599 GCAGCGGCTTCTCGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTTCTGG 658

RESULT 14  
 AL560880  
 LOCUS 1201 bp mRNA linear EST 31-MAY-2003  
 DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 ACCESION AL560880  
 VERSION Homo sapiens cDNA clone CSODL005YK19 5-PRIME, mRNA sequence.  
 KEYWORDS

AL560880.2 GI:31285009  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1201)  
 REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 AUTHORS Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12907768.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 7987.r For  
 more information about this cluster, see  
[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODL005AF100P1&cluster=7987.r)  
 Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CSODL005AF100P1.  
 Location/Qualifiers  
 1. 1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODL005YK19"  
 /cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 322 a 267 c 292 g 283 t 37 others  
 ORIGIN

Query Match 100.0%; Score 300; DB 9; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 60  
 Db 426 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 485  
 61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 120  
 Db 486 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGA 545  
 121 TGGAGATCAACACACAGCGGACTTTCCTCACACAAGCGCTCAACACATGTACAACTCC 180  
 Db 546 TGGAGATCAACACACAGCGGACTTTCCTCACACAAGCGCTCAACACATGTACAACTCC 605  
 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGAGCTCTAGAGAAAGGCTGTGT 240  
 Db 606 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGAGCTCTAGAGAAAGGCTGTGT 665  
 241 GCAGCGGCTTCTCGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTTCTGG 300  
 Db 666 GCAGCGGCTTCTCGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTTCTGG 725

RESULT 15  
 AL582250/c  
 LOCUS 1201 bp mRNA linear EST 01-JUN-2003  
 DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 ACCESION AL582250  
 VERSION Homo sapiens cDNA clone CSODL005YK11 3-PRIME, mRNA sequence.  
 KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT On Feb 16, 2001 this sequence version replaced gi:12950048.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7987.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DL005CC06NP1&cluster=7987.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DL005CC06NP1.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DL005YF11"  
/cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
/cell\_line="RAMOS CELL LINE"  
/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 291 a 238 c 281 g 293 t 98 others  
ORIGIN

## FEATURES

source

Query Match 100.0%; Score 300; DB 9; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60  
|||||  
Db 746 ACGAAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 687  
|||||

QY 61 CTGCTGACACTTGTGTAGACAGCAGGAGGCACATGCCAAGCTGGGTAACCTTGACCTTGA 120  
|||||  
686 CTGCTGACACTTGTGTAGACAGCAGGAGGCACATGCCAAGCTGGGTAACCTTGACCTTGA 627  
|||||

121 TGGAGATCAACACACAGCGGACTTTCCTCACAAAGGCTCAACCATGTACAAACTCC 180  
|||||

Db 626 TGGAGATCAACACACAGCGGACTTTCCTCACAAAGGCTCAACCATGTACAAACTCC 567  
|||||

QY 181 GCACGAACCTCCAGCCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
|||||

Db 566 GCACGAACCTCCAGCCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 507  
|||||

QY 241 GCAGGGCGCTTGTGGGGGATGTAGCGGCTCAGGACGTATGAGGTACTCGTGGTCTTGG 300  
|||||

Db 506 GCAGGGCGCTTGTGGGGGATGTAGCGGCTCAGGACGTATGAGGTACTCGTGGTCTTGG 447  
|||||

Search completed: October 14, 2003, 14:09:03  
Job time : 1279.67 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 12:04:45 ; Search time 157.333 Seconds  
(without alignments)  
5147.234 Million cell updates/sec

Title: US-09-854-124-5  
Perfect score: 300  
Sequence: 1 acgaataccgaccctggtc.....tgaagtactcgtgttctgg 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_15Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	ABQ78750	Differentially exp
2	300	100.0	858	AAH14111	Human cDNA sequenc
3	300	100.0	1171	ACAS4679	Human NF-kappaB as
4	300	100.0	1210	RAC77820	Human cancer assoc
5	292.4	97.5	667	AAH05624	Human cDNA clone (
6	252.8	84.3	1190	AAH87106	DNA encoding novel
7	252.8	84.3	1499	AAH87104	DNA encoding novel
8	240	80.0	678	AAH87103	DNA encoding novel

9	240	80.0	678	23	AAH87105	DNA encoding novel
10	214.4	71.5	300	24	ABQ78752	Differentially exp
11	140	46.7	454	24	ABL68902	Kidney cancer rela
12	100.6	33.5	603	24	ABQ59992	Human colon cancer
13	77.6	25.9	1257	23	AAH87051	DNA encoding novel
c 14	66	22.0	362	24	ABV88973	Human colon cancer
c 15	66	22.0	362	24	ABK45131	CDNA encoding colo
16	58.8	19.6	649	24	ABQ59972	Human colon cancer
17	53.8	17.9	612	23	ABL19607	Drosophila melanog
18	53.8	17.9	2670	23	ABL19606	Drosophila melanog
19	53.8	17.9	4504	23	ABL05636	Drosophila melanog
c 20	49.4	16.5	51	22	AAH27661	Human SNP oligonuc
21	39.6	13.2	734	22	AAH08244	Human cDNA clone (
22	39.6	13.2	940	23	AAH72563	DNA encoding novel
23	39.6	13.2	1143	25	ABX71197	Novel human cDNA s
24	39.6	13.2	3071	22	AAH16628	Human cDNA sequenc
25	36	12.0	654	24	ABN61061	Human cancer relat
26	35	11.7	2241	19	AAV68070	Maize-optimised DN
27	35	11.7	2370	19	AAV68069	Maize-optimised DN
28	35	11.7	2403	17	AAH13955	Maize-optimised VI
29	35	11.7	2403	18	AAH74004	Maize-optimised-B.
30	35	11.7	2403	19	AAV68065	Maize-optimised DN
31	35	11.7	2403	19	AAV16177	Maize-optimised DN
32	35	11.7	2444	17	AAH13964	VIP3A(a) synthetic
33	35	11.7	2444	18	AAH74011	B. cereus VIP3A(a)
34	35	11.7	2444	19	AAV16190	DNA for maize opti
35	33.8	11.3	360	21	AAZ50397	Human interferon-1
36	33.2	11.1	1206	24	ABN81296	Ancyclobacter form
37	33.2	11.1	1363	24	ABN81295	Ancyclobacter form
38	32.8	10.9	468	22	ABA42112	Human breast cell
39	32.8	10.9	468	22	ABA52534	Human foetal liver
40	32.8	10.9	468	22	ABA22323	Probe #789 for gen
41	32.8	10.9	468	22	AAK00797	Human brain expres
42	32.8	10.9	468	22	AAK26250	Human bone marrow
43	32.8	10.9	468	22	AAH10883	Probe #816 for gen
44	32.8	10.9	468	22	AAH32143	Probe #829 used to
45	32.8	10.9	468	22	AAH08006	Probe #797 used to

ALIGNMENTS

RESULT 1  
ABQ78750  
ID ABQ78750 standard; cDNA; 300 BP.  
XX AC ABQ78750;  
XX DT 05-DEC-2002 (first entry)  
XX DE Differentially expressed gene RTA00000684F.e.07.1.  
XX KW Differentially expressed gene; cancer; breast cancer; colon cancer;  
KW lung cancer; prostate cancer; expressed sequence tag; EST;  
KW heat shock factor binding protein; CGI-122 gene; quilescin Q6;  
KW moderately-differentiated endometrial adenocarcinoma; TGIF protein;  
KW treacher collins syndrome protein; annexin IV; cyclophilin C;  
KW MHC class I lymphocyte antigen; HLA-E class I mRNA;  
KW glomerulosclerosis gene; 265 proteasome subunit p55;  
KW gamma interferon inducible protein; cyclin-dependent protein kinase;  
KW mitochondrial dodecenoyl-CoA delta-isomerase;  
KW serine hydroxymethyltransferase;  
KW DNA-damage-inducible RNA binding protein; ss.  
OS Homo sapiens.  
XX US2002076735-A1.  
XX 20-JUN-2002.  
XX PF 10-MAY-2001; 2001US-0854124.  
XX PR 25-SEP-1998; 98US-101900P.



CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX  
 SQ Sequence 858 BP; 227 A; 214 C; 222 G; 195 T; 0 other;

Query Match 100.0%; Score 300; DB 22; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-86;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAATCCGGACCTGCTCAAGGATATGTGGGACACCTCGTATAGCCAAACTCCGAGTGT 60  
 Db |||||||  
 QY 61 CTGCTGCACAGCTTCTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACTTCA 120  
 Db |||||||  
 QY 428 CTGCTGCACAGCTTCTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACTTCA 487  
 Db |||||||  
 QY 121 TGGAGATCAACACACCGCGGACTTCTCTCACACAAGCGCTCAACACCATGTACAAACTCC 180  
 Db |||||||  
 QY 488 TGGAGATCAACACACCGCGGACTTCTCTCACACAAGCGCTCAACACCATGTACAAACTCC 547  
 Db |||||||  
 QY 181 GCACCAACTCTCAGCTCTGGAGAGTACTCAGTCTCAGGACCTCTAGAGAAAGCGCTGT 240  
 Db |||||||  
 QY 548 GCACCAACTCTCAGCTCTGGAGAGTACTCAGTCTCAGGACCTCTAGAGAAAGCGCTGT 607  
 Db |||||||  
 QY 241 GCAGCGGGCTTCTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 300  
 Db |||||||  
 QY 608 GCAGCGGGCTTCTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 667  
 Db |||||||

## RESULT 3

ACA54679  
 ID ACA54679 standard; cDNA; 1171 BP.

XX ACA54679;

XX  
 DT 05-JUN-2003 (first entry)

XX Human NF-kappaB associated polynucleotide sequence #119.

XX Human; nuclear factor-kappaB; NF-kappaB; immune disorder; cancer;  
 KW inflammatory disorder; apoptosis; hepatic disorder; Hodgkin's lymphoma;  
 KW haematopoietic tumour; hyper-IgM syndrome; viral infection; asthma;  
 KW hypohidrotic ectodermal dysplasia; human immunodeficiency virus; HIV;  
 KW X-linked anhidrotic ectodermal dysplasia; al incontinentia pigmenti;  
 KW influenza; rheumatoid arthritis; inflammatory bowel disease; colitis;  
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 KW experimental allergic encephalomyelitis; autoimmune disorder; wound;  
 KW hyper immune activity; acute phase response; hypercongenital condition;  
 KW birth defect; necrotic lesion; organ transplant rejection; pancreas;  
 KW signal transduction; hyperproliferative disorder; diabetes mellitus;  
 KW vitamin B12 malabsorption; neurological disorder; Huntington's chorea;  
 KW Turner's syndrome; bacterial infection; cardiovascular disorder;  
 KW infertility; psoriasis; haemolytic anaemia; antinflammatory; anti-HIV;  
 KW cytostatic; hepatotropic; viricide; antirheumatic; antiarthritic;  
 KW antiasthmatic; immunomodulator; antidiabetic; antiallergic;  
 KW neuroprotective; immunosuppressive; vulnerary; antibacterial;  
 KW antinfertility; antianemic; antipsoriatic; cerebroprotective;  
 KW cardiant; antiarteriosclerotic; gene; ss.

XX Homo sapiens.

OS WO200286076-A2.

XX  
 PN 31-OCT-2002.

XX 19-APR-2002; 2002WO-US12636.

XX 19-APR-2001; 2001US-284962P.

PR 26-APR-2001; 2001US-286645P.

PR 09-JAN-2002; 2002US-346986P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Carman J, Feder J, Nadler S;

XX WPI; 2003-093119/08.

DR P-PSDB; ABU69618.

XX Novel NF-kappaB-associated polypeptides and polynucleotides useful for  
 PT diagnosing, treating and preventing cancer, hepatic disorders, aberrant  
 PT apoptosis, viral infections, autoimmune disorders, asthma and stroke -  
 PS Claim 4; Fig 16; 608pp; English.

XX The present invention relates to the isolation of human nuclear  
 CC factor-kappaB (NF-kappaB) associated polypeptides and polynucleotides.  
 CC The NF-kappaB associated polypeptide and polynucleotide sequences  
 CC are useful for preventing, treating or ameliorating various disorders  
 CC including immune disorders, inflammatory disorders, cancers,  
 CC disorders relating to aberrant apoptosis, hepatic disorders,  
 CC Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM syndromes,  
 CC hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal  
 CC dysplasia, immunodeficiency, al incontinentia pigmenti, viral  
 CC infections (e.g. those caused by human immunodeficiency virus (HIV),  
 CC human T-cell lymphotropic virus (HTLV), hepatitis B, hepatitis C,  
 CC Epstein Barr virus (EBV), influenza), rheumatoid arthritis,  
 CC inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia,  
 CC euthyroid sick syndrome, stroke, experimental allergic encephalomyelitis,  
 CC (EAE), autoimmune disorders, disorders related to hyper immune activity,  
 CC disorders related to aberrant acute phase responses, hypercongenital  
 CC conditions, birth defects, necrotic lesions, wounds, organ transplant  
 CC rejection, disorders related to aberrant signal transduction,  
 CC hyperproliferative disorders, diseases of the pancreas (e.g. diabetes  
 CC mellitus, vitamin B12 malabsorption), neurological disorders (e.g.  
 CC Huntington's chorea), Turner's syndrome, bacterial infections,  
 CC cardiovascular disorders, infertility, psoriasis and haemolytic anaemia.  
 CC The present sequence represents a human NF-kappaB associated  
 CC polynucleotide of the invention.

XX Sequence 1171 BP; 296 A; 298 C; 309 G; 268 T; 0 other;

Query Match 100.0%; Score 300; DB 25; Length 1171;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-85;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAATCCGGACCTGCTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db |||||||  
 QY 408 ACGAATCCGGACCTGCTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 467  
 Db |||||||  
 QY 61 CTGCTGCACAGCTTCTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACTTCA 120  
 Db |||||||  
 QY 468 CTGCTGCACAGCTTCTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACTTCA 527  
 Db |||||||  
 QY 121 TGGAGATCAACACACCGCGGACTTCTCTCACACAAGCGCTCAACACCATGTACAAACTCC 180  
 Db |||||||  
 QY 528 TGGAGATCAACACACCGCGGACTTCTCTCACACAAGCGCTCAACACCATGTACAAACTCC 587  
 Db |||||||  
 QY 181 GCACCAACTCTCAGCTCTGGAGAGTACTCAGTCTCAGGACCTCTAGAGAAAGCGCTGT 240  
 Db |||||||  
 QY 588 GCACCAACTCTCAGCTCTGGAGAGTACTCAGTCTCAGGACCTCTAGAGAAAGCGCTGT 647  
 Db |||||||  
 QY 241 GCAGCGGGCTTCTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 300  
 Db |||||||  
 QY 648 GCAGCGGGCTTCTCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 707  
 Db |||||||

## RESULT 4

AAC77820

ID AAC77820 standard; cDNA; 1210 BP.

XX AAC77820;

XX

08-FEB-2001 (first entry)  
 Human cancer associated gene sequence SEQ ID NO:214.  
 Human; cancer associated gene; cancer antigen; detection; cancer; diagnosis; cytostatic; proliferative; vulnary; immunomodulator; antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral; antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant; dermatological; neuroprotective; thrombolytic; coagulant; nootropic; vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation; immune disorder; haematopoietic cell disorder; autoimmune disorder; allergic reaction; graft versus host disease; organ rejection; haemostatic; thrombolytic; cardiovascular disorder; infection; neurological disease; drug screening; ss.  
 Homo sapiens.  
 WO200055350-A1.  
 21-SEP-2000.  
 08-MAR-2000; 2000WO-US05882.  
 12-MAR-1999; 99US-0124270.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Ruben SM;  
 WPI; 2000-587533/55.  
 P-PSDB; AAB43611.  
 Novel isolated nucleic acids comprising sequences encoding peptides useful for treating or diagnosing e.g. cancer -  
 Claim 1; Page 785-786; 2352pp; English.  
 AAC77607 to AAC78448 encode the human cancer associated proteins given in AAB43398 to AAB44239. The proteins can have activities based on the tissues and cells the genes are expressed in. Example of activities include: cytostatic; proliferative; vulnary; immunomodulator; antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral; dermatological; neuroprotective; cardiant; thrombolytic; coagulant; nootropic; vasotropic; antipsoriatic and antiangiogenic. The polynucleotides and polypeptides can be used for preventing, treating or ameliorating medical conditions and diagnosing pathological conditions. Polynucleotides, polypeptides, antibodies, agonists and antagonists from the present invention may be used to treat immune disorders by activating or inhibiting the proliferation, differentiation or mobilisation of immune cells, to treat disorders of haematopoietic cells, autoimmune disorders, allergic reactions, graft versus host disease and organ rejection, modulate haemostatic or thrombolytic activity, modulate inflammation, cancers, cardiovascular disorders, neurological disease and bacterial or viral infections. The peptides, nucleotides, antibodies, agonists and antagonists may be also be used in drug screens. AAC78449 to AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.  
 Sequence 1210 BP; 320 A; 300 C; 320 G; 268 T; 2 other;  
 Query Match 100.0%; Score 300; DB 21; Length 1210;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-85;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ACAGAAATCCGACCTGCTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGACTGT 60  
 412 ACAGAAATCCGACCTGCTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGACTGT 471  
 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGACCTTGA 120  
 472 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATTAACCTTGACCTTGA 531

QY 121 TGGAGATCAACACACACGCGGAGCTTTCTTCCACACAAAGCGCTCAACACATGTACAAACTCC 180  
 DB 532 TGGAGATCAACACACACGCGGAGCTTTCTTCCACACAAAGCGCTCAACACATGTACAAACTCC 591  
 QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGGT 240  
 DB 592 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCTGGT 651  
 QY 241 GCAGGCGCTTGTCTGGGGATGTGAGCGCTCAGGAGCTGATGAGGACTCTCGTGGTCTGG 300  
 DB 652 GCAGGCGCTTGTCTGGGGATGTGAGCGCTCAGGAGCTGATGAGGACTCTCGTGGTCTGG 711

RESULT 5  
 AAH05624  
 ID AAH05624 standard; cDNA; 667 BP.  
 AC AAH05624;  
 DT 26-JUN-2001 (first-entry)  
 DE Human cDNA clone (5'-primer) SEQ ID NO:2459.  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX Homo sapiens.  
 XX EPI074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241899.  
 (HELI-) HELIX RES INST.  
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 WPI; 2001-318749/34.  
 Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -  
 Claim 1; SEQ ID 2459; 2537pp + CD ROM; English.  
 The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 667 BP; 171 A; 176 C; 177 G; 137 T; 6 other;

XX Query Match 97.5%; Score 292.4; DB 22; Length 667;

XX Best Local Similarity 97.7%; Pred. No. 2.3e-83; Indels 0; Gaps 0;

XX Matches 293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60

DB 367 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 426

QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATAACTTGCCTTGA 120

DB 427 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATAACTTGCCTTGA 486

QY 121 TGGAGATCAACACACCGGCGCTTTCCTCACACAGCGCTCAACCATGTACAACTCC 180

DB 487 TGGAGATCAACACACCGGCGCTTTCCTCACACAGCGCTCAACCATGTACAACTCC 546

QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCGCTGT 240

DB 547 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTANANAAGCGCTGT 606

QY 241 GCAGCGCGTCTCGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300

DB 607 GCAGCGCGTCTCGGGGATGTGAGCGCTCAGGAGCTGATNAAGTACTCGTGGTCTGG 666

#### RESULT 6

AAS87106

ID AAS87106 standard; cDNA; 1190 BP.

XX AC AAS87106;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22910.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABC22919.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 1; SEQ ID NO 22910; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1190 BP; 298 A; 304 C; 310 G; 278 T; 0 other;

XX Query Match 84.3%; Score 252.8; DB 23; Length 1190;

XX Best Local Similarity 98.0%; Pred. No. 1.4e-70;

XX Matches 298; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60

DB 413 ACGAATCCGGACCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 472

QY 61 CTGCTGACAGC-TTGTGTGACAGCAGGAGGACATGCCAAGCT-CGATAACTTGACCTT 118

DB 473 CTGCTGACAGC-TTGTGTGACAGCAGGAGGACATGCCAAGCTCGGTAACCTTGACTT 532

QY 119 -GATGGAGATCAACACCGCGG-GACTTTCCTCACACAGCGCTCAACCATGTACAAA 176

DB 533 CGATGGAGATCCACACCGCGGTGACTTTCCCTCACACAGCGCTCAACCATGTACAAA 592

QY 177 CTCGCCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCC 236

DB 593 CTCGCCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCC 652

QY 237 TGGTGCAGCGCGCTTGTGAGCGCTCAGGAGCTGAGCGCTCAGGAGTACTCGTGGTT 296

DB 653 TGGTGCAGCGCGCTTGTGAGCGCTCAGGAGTACTCGTGGAGTACTCGTGGTT 712

QY 297 CTGG 300

DB 713 CTGG 716

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABC22919.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 1; SEQ ID NO 22910; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

#### RESULT 7

AAS87104

ID AAS87104 standard; cDNA; 1499 BP.

XX AC AAS87104;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22908.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.



Db 580 GCACAACTCAGCCTTCGGAGAGTACTCACTCAGGACTTCTAGAGAAAGGCGCTGCT 639

QY 241 GCAGCGCGCTTCTCGGGGATGTGACGCTCAGGACGTGA 280

Db 640 GCAGGACGCTTCTCGGGGATGTGA-CGCTCAGGACGTGA 678

# RESULT 9

AA587105

ID AAS87105 standard; cDNA; 678 BP.

AC AAS87105;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22909.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR P-PSDB; ABG22918.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

XX Claim 1; SEQ ID No 22909; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 678 BP; 177 A; 185 C; 206 G; 110 T; 0 other;

Query Match 80.0%; Score 240; DB 23; Length 678;

Best Local Similarity 95.7%; Pred. No. 1.4e-66;

Matches 268; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 ACGAAATCGGACCTGTGTCAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
Db 401 ACGAAATCGGACCTGTGTCAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 460  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATTAACCTTGA 120  
Db 461 CTGCTGACAGCTTTGTGAGGCAACAGGAGGACATGCCAAGCTGGATTAACCTTGA 520  
QY 121 TGGAGATCAACACACAGCGGACTTTCTCACAACAAGCGCTCAACACATGTACAACTCC 180  
Db 521 TGGAGATCAACCCAGCGGACTTTCTCACAACAAGCGCTCAACACATGTACAACT-C 579  
QY 181 GCACGAACCTCAGCCTCTGGAGAGTACTCAGCTCAGGACTTCTAGAGAAAGGCGCTGT 240  
Db 580 GCACAAACCTCAGCCTCTGGAGAGTACTCAGCTCAGGACTTCTAGAGAAAGGCGCTGT 639  
QY 241 GCAGCGCGCTTCTCGGGGATGTGACGCTCAGGACGTGA 280  
Db 640 GCAGGACGCTTCTCGGGGATGTGA-CGCTCAGGACGTGA 678

## RESULT 10

ABQ78752

ID ABQ78752 standard; cDNA; 300 BP.

XX ABQ78752;

XX 05-DEC-2002 (first entry)

DE Differentially expressed gene RTA00000596f.d.12.1.

XX Differentially expressed gene; cancer; breast cancer; colon cancer;  
KW lung cancer; prostate cancer; expressed sequence tag; EST;  
KW heat shock factor binding protein; CGI-122 gene; quiescin Q6;  
KW moderately-differentiated endometrial adenocarcinoma; TGIF protein;  
KW treacher collins syndrome protein; annexin IV; cyclophilin C;  
KW MHC class I lymphocyte antigen; HLA-E class I MRNA;  
KW glomerulosclerosis gene; 265 proteasome subunit p55;  
KW gamma interferon inducible protein; cyclin-dependent protein kinase;  
KW mitochondrial dodecenoyl-CoA delta-isomerase;  
KW serine hydroxymethyltransferase;  
KW DNA-damage-inducible RNA binding protein; ss.

XX Homo sapiens.

OS US2002076735-A1.

XX 20-JUN-2002.

XX 10-MAY-2001; 2001US-0854124.

XX 25-SEP-1998; 98US-101900P.

XX 22-SEP-1999; 99US-0400947.

PA (WILL/) WILLIAMS L T.

PA (ESCO/) ESCOBEDO J.

PA (INNI/) INNIS M A.

PA (GARC/) GARCIA P D.

PA (SUDD/) SUDDUTH-KLINGER J.

PA (REIN/) REINHARD C.

PA (GIES/) GIESE K.

PA (RAND/) RANDAZZO F.

PA (KENN/) KENNEDY G C.

PA (POTD/) POT D.

PA (KASS/) KASSAM A.

PA (LAWS/) LAWS G.

PA (DRMA/) DRMANAC R.

PA (CRKV/) CRKVENJAKOV R.

PA (DICK/) DICKSON M.

PA (DRMA/) DRMANAC S.

PA (LABA/) LABAT I.

PA (LESH/) LESHKOWITZ D.

PA (KITA/) KITA D.





XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664  
CC to AB170110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 454 BP; 128 A; 113 C; 116 G; 97 T; 0 other;

Query Match 46.7%; Score 140; DB 24; Length 454;  
Best Local Similarity 99.3%; Pred. No. 1.2e-34;  
Matches 151; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 60  
DB 303 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 362  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACCTTCA 120  
DB 363 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACCTTCA 421  
QY 121 TGGAGATCAACACACCGCGGACTTTCTCTCAC 152  
DB 422 TGGAGATCAACACACCGCGGACTTTCTCTCAC 453

RESULT 12

ABQ59992  
ID ABQ59992 standard; cdna; 603 BP.

AC ABQ59992;

DT 02-AUG-2002 (first entry)

DE Human colon cancer related nucleotide sequence SEQ ID NO:3687.

KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;

KW genetic analysis; diagnostic; antisense therapy; gene; ss.

OS Homo sapiens.

PN WO200229086-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001WO-US30732.

PR 02-OCT-2000; 2000US-237271P.

PA (FARB ) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
PI Thianglingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell  
PT or tissue type, and in antisense therapy -

PS Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins  
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence of  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists.

XX Sequence 603 BP; 157 A; 161 C; 162 G; 117 T; 6 other;

Query Match 33.5%; Score 100.6; DB 24; Length 603;  
Best Local Similarity 89.0%; Pred. No. 5.4e-22;  
Matches 154; Conservative 0; Mismatches 11; Indels 8; Gaps 4;

QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 60  
DB 394 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 452  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACCTT 118  
DB 453 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACCTT 512  
QY 119 GAT-GGAGATCAACACCA---CGCGGACTTTCTCTCACACACGCGCTCAACCA 166  
DB 513 GATGGAGATCAACACCAAGCAGCGGACTTTCTCTCACACACGCGCTCAACCA 565

RESULT 13

AAS83051  
ID AAS83051 standard; cdna; 1257 BP.

AC AAS83051;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #18855.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG18864.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

XX

Claim 1: SEQ ID No 18855; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polynucleotide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

[illegible]

RESULT 14	
ABV88973/c	
ID	ABV88973 standard; cDNA; 362 BP.
XX	
AC	ABV88973;
XX	
XX	DT 13-DEC-2002 (first entry)
XX	
DE	Human colon cancer related cDNA SEQ ID NO 2284.
DE	
DE	Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
KW	ss.
KW	
XX	
XX	
OS	Homo sapiens.
XX	
PN	WO200258534-A2.
XX	
PD	01-AUG-2002.
XX	
PF	19-NOV-2001; 2001WO-US43704.
XX	
PR	20-NOV-2000; 2000US-252222P.
XX	

PR	06-FEB-2001; 2001US-267011P.
PR	28-MAR-2001; 2001US-279670P.
PR	10-JUL-2001; 2001US-304037P.
XX	(CORI-) CORIXA CORP.
XX	Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX	WPI; 2002-608400/65.
DR	New isolated tumor colon polynucleotide and polypeptide, useful for the
XX	diagnosis, prevention and/or treatment of cancer, in particular colon
PT	cancer -
PT	
XX	ClaIm 1; SEQ ID NO 2284; 266pp + Sequence Listing; English.
XX	
PS	The invention relates to a human colon tumour expressed polynucleotide
CC	(I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
CC	2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC	complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
CC	sequences that hybridize to (i), under moderately stringent conditions;
CC	(v) sequences having at least 75% or 90% identity to (i); or (vi)
CC	degenerate variants of (i). The compositions and methods of the present
CC	invention are useful for the diagnosis, prevention and/or treatment of
CC	cancer, particularly colon cancer. (I) can be used in gene therapy and
CC	(I) and (II) are useful in pharmaceutical compositions such as vaccines.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at <a href="http://wipo.int/pub/published.pct.sequences">ftp.wipo.int/pub/published.pct.sequences</a> .

```

SQ Sequence 362 BP; 100 A; 98 C; 73 G; 91 T; 0 other;
      22.0%; Score 66; DB 24; Length 362;
Query Match
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 235 CCTGGTGCAGCGCGCTTGCTGGGGGATGTGAGCGCTCAGGACGFTGATGAGGTACTCTCGTGG 294
    |||||
Db 362 CCTGGTGCAGCGCGCTTGCTGGGGGATGTGAGCGCTCAGGACGFTGATGAGGTACTCTCGTGG 303
    |||||
QY 295 TTCTGG 300
    |||||
Db 302 TTCTGG 297

```

RESULT 15  
ABR45131/C  
ID ABR45131 standard; cDNA; 362 BP.  
XX  
XX ABR45131;  
AC  
XX  
XX  
XX  
DT  
XX  
DE 05-JUN-2002 (first entry)  
DE cDNA encoding colon tumour protein, SEQ ID No 682.  
XX  
XX  
XX Human; colon tumour; vaccine; colon cancer; immunogenic;  
KW immunotherapy; gene; ss.  
KW  
XX  
XX  
OS Homo sapiens.  
XX  
XX WO200212328-A2.  
XX  
PD  
PD 14-FEB-2002.  
XX  
XX  
PF 31-JUL-2001; 2001WO-US24218.  
XX  
XX 03-AUG-2000; 2000US-223283P.  
PR 28-MAR-2001; 2001US-279763P.  
PR 29-JUN-2001; 2001US-302051P.  
XX  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX King GE. Meagher MJ. Xu J. Secret H:

```

XX WPI: 2002-241739/29.
XX
PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers
PT for the progression of cancer -
XX
PS Claim 1: SEQ ID No 682; 147pp; English.
XX
CC The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK44450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.
CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.
XX
SQ Sequence 362 BP; 99 A; 98 C; 74 G; 91 T; 0 other;

Query Match      22.0%; Score 66; DB 24; Length 362;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CCTGGTCAGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 294
Db 362 CCTGGTCAGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGG 303
QY 295 TTCTGG 300
Db 302 TTCTGG 297

Search completed: October 14, 2003, 14:17:01
Job time : 159.333 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 12:04:45 ; Search time 157.333 Seconds  
(without alignments)  
5147.234 Million cell updates/sec

Title: US-09-854-124-6

Perfect score: 300

Sequence: 1 aattccgtgtcgtcgtga.....cgaaagaagaacttttac 300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_15Jun03.\*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	24: ABQ78751	Differentially exp
2	284.8	94.9	603	24: ABQ59992	Human colon cancer
3	284.8	94.9	858	22: AAH14111	Human CDNA sequenc
4	284.8	94.9	1171	25: ACA54679	Human NF-kappaB as
5	284.8	94.9	1210	21: AAC77820	Human cancer assoc
6	283	94.3	667	22: AAH05624	Human CDNA clone (
7	281.2	93.7	649	24: ABQ59972	Human colon cancer
8	273.8	91.3	1190	23: AAS87106	DNA encoding novel

9	231.4	77.1	1257	23	AAS83051	DNA encoding novel
10	223.4	74.5	454	24	ABL68902	Kidney cancer rela
11	102.8	34.3	678	23	AAS87103	DNA encoding novel
12	102.8	34.3	678	23	AAS87105	DNA encoding novel
13	90	30.0	1499	23	AAS87104	DNA encoding novel
14	85.2	28.4	612	23	ABL19607	Drosophila melanog
15	85.2	28.4	2870	23	ABL19606	Drosophila melanog
16	85.2	28.4	4904	23	ABL03636	Drosophila melanog
17	46.8	15.6	525	25	ABZ53276	Aspergillus oryzae
18	37.2	12.4	29559	23	AAS59546	Propionibacterium
19	36	12.0	2697	24	AAS17590	DNA encoding novel
20	34.8	11.6	1018	16	AA02326	Human cardiac hype
21	34.8	11.6	1018	18	AA07014	Human cardiocroph
22	34.8	11.6	1539	20	AA087267	cDNA clone encodin
23	34.8	11.6	1539	21	AAA99914	cDNA encoding huma
24	34.8	11.6	1539	21	AAA72705	Human cardiocroph
25	34.8	11.6	1539	21	AAA72706	Human cardiocroph
26	34.8	11.6	1539	21	AAA46936	cDNA encoding nove
27	34.8	11.6	1539	22	AA090578	Human PRO882 cDNA
28	34.8	11.6	1539	24	ABL63697	Breast cancer rela
29	34.8	11.6	3161	24	AA049189	Porcine CD 151 gen
30	34.8	11.6	61313	23	AAS59545	Propionibacterium
31	34.6	11.5	1626	25	ABT31957	High-affinity phos
32	34.6	11.5	1626	25	ABT32039	High-affinity OSPH
33	34.2	11.4	277	24	ABN23362	Human ORFX polynuc
34	33.8	11.3	3899	24	AB199205	Mouse ischaemic co
35	33.6	11.2	936	24	AAS17241	Streptomyces clavu
36	33.6	11.2	3773	23	ABL11845	Drosophila melanog
37	33.6	11.2	17951	23	ABL11844	Drosophila melanog
38	33.4	11.1	1118	24	ABK75437	Bacillus lichenifo
39	33	11.0	2000	24	AB217573	Arabidopsis thalia
40	33	11.0	2577	24	AA041592	HIV-1 subtype C is
41	33	11.0	3981	24	AA167898	Nucleotide sequenc
42	33	11.0	4288	22	AAH20868	HIV-1 subtype (C/B
43	33	11.0	4341	24	AA167897	Nucleotide sequenc
44	33	11.0	4343	24	AA167896	Nucleotide sequenc
45	32.6	10.9	480	20	AA001141	M. tuberculosis an

#### ALIGNMENTS

##### RESULT 1

ABQ78751  
ID ABQ78751 standard; cDNA; 300 BP.

XX AC ABQ78751;

XX DT 05-DEC-2002 (first entry)

XX DE Differentially expressed gene RTA00000618F.p.24.1.

XX KW Differentially expressed gene; cancer; breast cancer; colon cancer;  
KW lung cancer; prostate cancer; expressed sequence tag; EST;  
KW heat shock factor binding protein; CGI-122 gene; quiescin Q6;  
KW moderately-differentiated endometrial adenocarcinoma; TGIF protein;  
KW treacher collins syndrome protein; annexin IV; cyclophilin C;  
KW MHC class I lymphocyte antigen; HLA-E class I mRNA;  
KW glomerulosclerosis gene; 265 proteasome subunit p55;  
KW gamma interferon inducible protein; cyclin-dependent protein kinase;  
KW mitochondrial dodecenoyl-CoA delta-isomerase;  
KW serine hydroxymethyltransferase;  
KW DNA-damage-inducible RNA binding protein; ss.

XX OS Homo sapiens.

XX PN US2002076735-A1.

XX PD 20-JUN-2002.

XX PF 10-MAY-2001; 2001US-0854124.

XX PR 25-SEP-1998; 98US-101900P.

Db	61	GGTCGAATTCTCGCGAAGAGCTGGTTACCATTTACCCCAACTTCAGTCTGCACAA	120
QY	121	GATCTACCTCATCGGGGGGACCTGGGGCCCTTTTAACTCCCTGGTTTACCCGTGGAATGCC	180
Db	121	GATCTACCTCATCGGGGGGACCTGGGGCCCTTTTAACTCCCTGGTTTACCCGTGGAATGCC	180
QY	181	CTGTGGCTGGCGATTAACTGAAACAAAGACAGAAATGTCGCTCTCCTCCCTCCAGATG	240
Db	181	CTGTGGCTGGCGATTAACTGAAACAAAGACAGAAATGTCGCTCTCCTCCCTCCAGATG	240
QY	241	GATGGATGTAGAAAACTTGGAGAAGATGAGGGATCATGAACGAAAGGAAGAACTTTTAC	300
Db	241	GATGGATGTAGAAAACTTGGAGAAGATGAGGGATCATGAACGAAAGGAAGAACTTTTAC	300
RESULT 2			
ABQ59992			
ID	ABQ59992 standard; cDNA; 603 BP.		
AC	ABQ59992;		
XX			
DT	02-AUG-2002 (first entry)		
XX	Human colon cancer related nucleotide sequence SEQ ID NO:3687.		
XX			
DE			
XX			
KW	Human: colon cancer; cancer; tissue profiling; forensic; mapping;		
KW	genetic analysis; diagnostic; antisense therapy; gene; ss.		
XX			
OS	Homo sapiens.		
OS			
WO	WO200229086-A2.		
PN			
XX			
PD	11-APR-2002.		
XX			
XX	02-OCT-2001; 2001WO-US30732.		
PF			
XX			
PR	02-OCT-2000; 2000US-237271P.		
XX			
PA	(FARB ) BAYER CORP.		
XX			
PI	Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;		
PI	Thiagalingam A, Lewis ME;		
XX			
DR	WPI; 2002-426115/45.		
XX			
PT	New isolated nucleic acid that is differentially expressed in cancer		
PT	tissues useful for determining the presence of colon cancer in a cell		
PT	or tissue type, and in antisense therapy		
XX			
PS	Claim 1; Fig 1; 796pp; English.		
XX			
CC	ABQ59306 to ABQ60787 represent isolated nucleic acids (I) differentially		
CC	expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins		
CC	encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be		
CC	used in antisense therapy. An antibody immunoreactive with a polypeptide		
CC	encoded by (I) is useful for detecting cancer in a patient sample, and		
CC	for detecting the presence or absence of a polynucleotide encoded by a		
CC	nucleic acid which hybridises to (I) in a cell. A probe/primer derived		
CC	from (I) can be used for determining the presence of a nucleic acid which		
CC	hybridises to (I), and for determining the phenotype of cells in a sample		
CC	of cells from a patient. (I) is useful for determining the presence of		
CC	colon cancer in a cell or tissue type, for determining the presence or		
CC	state of other type of cancer, in antisense therapy, to generate		
CC	macroarrays on a solid surface, to identify a chromosome on which the		
CC	corresponding gene resides, and in tissue profiling, forensics, genetic		
CC	analysis, mapping and diagnostic applications. (I) can be used to raise		
CC	antibodies, and to screen for peptide analogues and antagonists.		
XX			
SQ	Sequence 603 BP; 157 A; 161 C; 162 G; 117 T; 6 other;		

Query Match 94.9%; Score 284.8; DB 24; Length 603;

Best Local Similarity 99.3%; Pred. No. 9,3e-78;



XX 31-OCT-2002.  
 XX 19-APR-2002; 2002WO-US12636.  
 XX 19-APR-2001; 2001US-284962P.  
 XX 26-APR-2001; 2001US-286645P.  
 XX 09-JAN-2002; 2002US-346986P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX Carman J, Feder J, Nadler S;  
 XX WPI; 2003-093119/08.  
 XX P-PSDB; AB069618.  
 XX Novel NF-kappaB-associated polypeptides and polynucleotides useful for  
 XX diagnosing, treating and preventing cancer, hepatic disorders, aberrant  
 XX apoptosis, viral infections, autoimmune disorders, asthma and stroke -  
 XX Claim 4; Fig 16; 608pp; English.  
 XX The present invention relates to the isolation of human nuclear  
 XX factor-kappaB (NF-kappaB) associated polypeptides and polynucleotides.  
 XX The NF-kappaB associated polypeptide and polynucleotide sequences  
 XX are useful for preventing, treating or ameliorating various disorders  
 XX including immune disorders, inflammatory disorders, cancers,  
 XX disorders relating to aberrant apoptosis, hepatic disorders,  
 XX Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM syndromes,  
 XX hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal  
 XX dysplasia, immunodeficiency, al incontinentia pigmenti, viral  
 XX infections (e.g. those caused by human immunodeficiency virus (HIV),  
 XX human T-cell lymphotropic virus (HTLV), hepatitis B, hepatitis C,  
 XX Epstein Barr virus (EBV), influenza), rheumatoid arthritis,  
 XX inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia,  
 XX euthyroid sick syndrome, stroke, experimental allergic encephalomyelitis  
 XX (EAE), autoimmune disorders, disorders related to hyper immune activity,  
 XX disorders related to aberrant acute phase responses, hypercongenital  
 XX conditions, birth defects, necrotic lesions, wounds, organ transplant  
 XX rejection, disorders related to aberrant signal transduction,  
 XX hyperproliferative disorders, diseases of the pancreas (e.g. diabetes  
 XX mellitus, vitamin B12 malabsorption), neurological disorders (e.g.  
 XX Huntington's chorea), Turner's syndrome, bacterial infections, e.g.  
 XX cardiovascular disorders, infertility, psoriasis and haemolytic anaemia.  
 XX The present sequence represents a human NF-kappaB associated  
 XX polynucleotide of the invention.  
 XX  
 XX Sequence 1171 BP; 296 A; 298 C; 309 G; 268 T; 0 other;  
 XX  
 XX Query Match 94.9%; Score 284.8; DB 25; Length 1171;  
 XX Best Local Similarity 99.3%; Pred. No. 1.2e-77;  
 XX Mismatches 0; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 XX 13 GTCGGTGAGGCTGTGGCTGCGAGCTGCGCGCCGATGGAGCTGCGAGGTGCAATTCCT 72  
 XX 43 GACCGTGAGGCTGTGGCTGCGAGCTGCGCGCCGATGGAGCTGCGAGGTGCAATTCCT 102  
 XX 73 CGCCGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGCTGCGAAGATCTACCTCAT 132  
 XX 103 CGCCGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGCTGCGAAGATCTACCTCAT 162  
 XX 133 CGGGGGGACCTGGGGCCCTTTAACTGCTGGTTACCGTGGAGTGGCCCTGTGGCTGGC 192  
 XX 163 CGGGGGGACCTGGGGCCCTTTAACTGCTGGTTACCGTGGAGTGGCCCTGTGGCTGGC 222  
 XX 193 GATTACCTGAAACAAAGACAAATGTCGCTGCTCCCTCCAGAGTGATGATGTAGA 252  
 XX 223 GATTACCTGAAACAAAGACAAATGTCGCTGCTCCCTCCAGAGTGATGATGTAGA 282  
 XX 253 AAAGTTGGAGAGATGAGGATCATCAACGAAAGGAAGAACTTTTAC 300  
 XX 283 AAAGTTGGAGAGATGAGGATCATCAACGAAAGGAAGAACTTTTAC 330

RESULT 5  
 AAC77820  
 ID AAC77820 standard; cDNA; 1210 BP.  
 XX AAC77820;  
 AC AAC77820;  
 XX 08-FEB-2001 (first entry)  
 DT 08-FEB-2001 (first entry)  
 DE Human cancer associated gene sequence SEQ ID NO:214.  
 XX Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;  
 KW antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral;  
 KW antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; nontropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX WO200055350-A1.  
 PN WO200055350-A1.  
 XX 21-SEP-2000.  
 PD 21-SEP-2000.  
 XX 08-MAR-2000; 2000WO-US05882.  
 PF 08-MAR-2000; 2000WO-US05882.  
 XX 12-MAR-1999; 99US-0124270.  
 PR 12-MAR-1999; 99US-0124270.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM;  
 PI Rosen CA, Ruben SM;  
 XX WPI; 2000-587533/55.  
 DR WPI; 2000-587533/55.  
 DR P-PSDB; AAB43611.  
 XX Novel isolated nucleic acids comprising sequences encoding peptides  
 XX useful for treating or diagnosing e.g. cancer -  
 XX Claim 1; Page 785-786; 2352pp; English.  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 XX in AAB43398 to AAB44239. The proteins can have activities based on the  
 XX tissues and cells the genes are expressed in. Example of activities  
 XX include: cytostatic; proliferative; vulnary; immunomodulator;  
 XX antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral;  
 XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 XX nontropic; vasotropic; antipsoriatic and antiangiogenic. The  
 XX polynucleotides and polypeptides can be used for preventing, treating or  
 XX ameliorating medical conditions and diagnosing pathological conditions.  
 XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 XX the present invention may be used to treat immune disorders by activating  
 XX or inhibiting the proliferation, differentiation or mobilisation of  
 XX immune cells, to treat disorders of haematopoietic cells, autoimmune  
 XX disorders, allergic reactions, graft versus host disease and organ  
 XX rejection, modulate haemostatic or thrombolytic activity, modulate  
 XX inflammation, cancers, cardiovascular disorders, neurological disease and  
 XX bacterial or viral infections. The peptides, nucleotides, antibodies,  
 XX agonists and antagonists may be also be used in drug screens. AAC78449 to  
 XX AAC78457 and AAB44240 represent sequences used in the exemplification of  
 XX the present invention.  
 XX  
 XX Sequence 1210 BP; 320 A; 300 C; 320 G; 268 T; 2 other;  
 XX  
 XX Query Match 94.9%; Score 284.8; DB 21; Length 1210;  
 XX Best Local Similarity 99.3%; Pred. No. 1.2e-77;  
 XX Mismatches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 XX 13 GTCGGTGAGGCTGTGGCTGCGAGCTGCGCGCCGATGGAGCTGCGAGGTGCAATTCCT 72



|||||  
47 GACCTGAGGCTCTGGCCCTGCAGCTCGCCGCCCATGAGCGCTGCCGAGGTGCAATCTCT 106  
|||||  
73 CGCCGAGAAGGAGCTGGTTTACCATTTATCCCACTTACGCTCGGACAAGATCTACCTCAT 132  
|||||  
107 CGCCGAGAAGGAGCTGGTTTACCATTTATCCCACTTACGCTCGGACAAGATCTACCTCAT 166  
|||||  
133 CGGGGGGACCTGGGGGCTTTTAACTCGTGGTTTACCGTGGGAAGTGCCTGTGGCTGGC 192  
|||||  
167 CGGGGGGACCTGGGGGCTTTTAACTCGTGGTTTACCGTGGGAAGTGCCTGTGGCTGGC 226  
|||||  
193 GATTAACCTGAAACAAAGACAGAAATGTGCTGCTCCCTCCAGAGTGGATGATGTAGATA 252  
|||||  
227 GATTAACCTGAAACAAAGACAGAAATGTGCTGCTCCCTCCAGAGTGGATGATGTAGATA 286  
|||||  
253 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAAACTTTTAC 300  
|||||  
287 AAAGTTGGAGAAGATGAGGATCATGAACGAAGGAAGAAACTTTTAC 334  
|||||  
RESULT 6  
AAH05624  
ID AAH05624 standard; cDNA; 667 BP.  
XX  
AC AAH05624;  
XX  
XX 26-JUN-2001 (first entry)  
XX  
DE Human cDNA clone (5'-primer) SEQ ID NO:2459.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs -  
XX  
XX Claim 1; SEQ ID 2459; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification. Where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and

in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow obtaining of the full-length  
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.  
XX  
SQ Sequence 667 BP; 171 A; 176 C; 177 G; 137 T; 6 other;  
Query Match 94.3%; Score 283; DB 22; Length 667;  
Best Local Similarity 100.0%; Pred. No. 3.5e-77;  
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 18 TGAGGCTCTGGGCTCGCAGCTCGCGGCCCATGACGCTGCCGAGGTGCAATCTCTCGCG 77  
DB 7 TGAGGCTCTGGGCTCGCAGCTCGCGGCCCATGACGCTGCCGAGGTGCAATCTCTCGCG 66  
QY 78 AGAAGGAGCTGGTTTACCATTTATCCCACTTACGCTCGGACAAGATCTACCTCATCGGG 137  
DB 67 AGAAGGAGCTGGTTTACCATTTATCCCACTTACGCTCGGACAAGATCTACCTCATCGGG 126  
QY 138 GGGACCTGGGGCTTTTAACTCGTGGTTTACCGTGGGAAGTGCCTGTGGCGGATTA 197  
DB 127 GGGACCTGGGGCTTTTAACTCGTGGTTTACCGTGGGAAGTGCCTGTGGCGGATTA 186  
QY 198 ACCTGAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATAGAAAGT 257  
DB 187 ACCTGAACAAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGGATAGAAAGT 246  
QY 258 TGGAGAAGATGAGGGATCATGAACGAAGGAAGAAACTTTTAC 300  
DB 247 TGGAGAAGATGAGGGATCATGAACGAAGGAAGAAACTTTTAC 289  
RESULT 7  
ABQ59972  
ID ABQ59972 standard; cDNA; 649 BP.  
XX  
AC ABQ59972;  
XX  
DT 02-AUG-2002 (first entry)  
XX  
DE Human colon cancer related nucleotide sequence SEQ ID NO:3667.  
XX  
DE Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200229086-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US30732.  
XX  
PR 02-OCT-2000; 2000US-237271P.  
XX  
PA (FARB ) BAYER CORP.  
XX  
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
PI Thialingam A, Lewis ME;  
XX  
WPI; 2002-426115/45.  
XX  
XX New isolated nucleic-acid that is differentially expressed in cancer  
tissues useful for determining the presence of colon cancer in a cell  
or tissue type, and in antisense therapy -  
XX  
XX Claim 1; Fig 1; 796pp; English.  
XX



PN WO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 XX 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Dmanac RT, Liu C, Tang YT;  
 XX WPI: 2001-639362/73.  
 XX P-PSDB; ABG18864.  
 DR New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1: SEQ ID No 18855; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and  
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 1257 BP; 317 A; 311 C; 332 G; 296 T; 1 other;  
 SQ

Query Match  
 Best Local Similarity 77.1%; Score 231.4; DB 23; Length 1257;  
 Matches 280; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

Qy	13	GTCCGTGAGGCTTGCGCTCAGCTCGCGCCGCGCCGAGTGCCTGCGAGGTGCAATTCCT	72
Db	44	GACCGTGAGGCTTGCGCTCAGCTCGCGCCGCGCCGAGTGCCTGCGAGGTGCAATTCCT	103
Qy	73	CGCCGAGAAGGAGCTGTTACCATATATCCCACTTCAGTCTGCAGCAAGATCTACCTCAT	132
Db	104	CGCCGAGAAGGAGCTGTTACCATATATCCCACTTCAGTCTGCAGCAAGATCTACCTCAT	163
Qy	133	CGGGGGGAGCTGGGGCTTTTAAACCTT-GTTTTACCCGTGGAGTCCCTCTGGGCTGG	191
Db	164	CGGGGGGAGCTGGGGCTTTTAAACCTTGGTTTACCCGTGGAGTCCCTCTGGGCTGG	223
Qy	192	--CGATTAACTGAAACAAAGACAG-AAAATGTCGCTGCTCCCTCCAGAGTGGATGGATG	248
Db	224	CGGATTAACTGAAACAAAGACAGAAAATGTCGCTGCTCCCTCCAGAGTGGATGGATG	283
Qy	249	TAGAAAAGTTCGAGAAATGAGGATCATGAA-CGAAAGGAAGAACTTTT	298
Db	284	TAGAAAAGTTCGAGAAATGAGGATCATGAAACCGAAAGGAAGAACTTTT	334

RESULT 10

ABL68902  
 ID ABL68902 standard; DNA; 454 BP.  
 XX  
 AC ABL68902;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Kidney cancer related gene sequence SEQ ID NO:7239.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI; 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a





OS Drosophila melanogaster.  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 10294; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 612 BP; 155 A; 166 C; 138 G; 138 T; 0 other;

Query Match 28.4%; Score 85.2; DB 23; Length 612;  
 Best Local Similarity 60.3%; Pred. No. 3.8e-16;  
 Matches 141; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
 QY 47 ATGGACGCTCGGAGGTGCAATTCCTCGCGAGAAGGAGCTGTACCATTTATCCCAAC 106  
 DB 1 ATGGATCCTTCAATTTATGATTTATTTGGCGAAAATGCATGATCAGCATAAATACCGAAC 60  
 QY 107 TTCAGTCTGGACAAGATCTACCTCATCGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTA 166  
 DB 61 TTCAGCAACGAGCCTCTGCACCTGATATACGACCTGTGGGCCCTTTCCGAGCGGTTT 120  
 QY 167 CCCGTGGAAGTGGCCCTGTGGCTGGCGATTAACCTGAAACAAGACAGAAATGTCGCTG 226  
 DB 121 CCCGTCTTGGTGGCCCTGTGGATGGCCACGCTGTGCGCAAGCAACAAAAGTCCGGAAT 180  
 QY 227 CTCCTCCAGATGGATGGATGTAGAAAAGTTGGAGAAGATGAGGGATCATGAA 280  
 DB 181 GTACCTCCAGATGGATGGATGATATATTGGAGGAATCAAGGAGGAGAA 234

RESULT 15  
 ABL19606  
 ID ABL19606 standard; DNA; 2670 BP.  
 XX  
 AC ABL19606;  
 XX

26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 10291.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical; gene; ds.  
 XX  
 OS Drosophila melanogaster.  
 XX

PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1; SEQ ID NO 10291; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2670 BP; 733 A; 627 C; 635 G; 675 T; 0 other;

Query Match 28.4%; Score 85.2; DB 23; Length 2670;  
 Best Local Similarity 60.3%; Pred. No. 7e-16;  
 Matches 141; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
 QY 47 ATGGACGCTCGGAGGTGCAATTCCTCGCGAGAAGGAGCTGTACCATTTATCCCAAC 106  
 DB 1001 ATGGATCCTTCAATTTATGATTTATTTGGCGAAAATGCATGATCAGCATAAATACCGAAC 1060  
 QY 107 TTCAGTCTGGACAAGATCTACCTCATCGGGGGGACCTGGGGCCCTTTTAAACCCCTGGTTTA 166  
 DB 1061 TTCAGCAACGAGCCTCTGCACCTGATATACGACCTGTGGGCCCTTTCCGAGCGGTTT 1120  
 QY 167 CCCGTGGAAGTGGCCCTGTGGCTGGCGATTAACCTGAAACAAGACAGAAATGTCGCTG 226  
 DB 1121 CCCGTCTTGGTGGCCCTGTGGATGGCCACGCTGTGCGCAAGCAACAAAAGTCCGGAAT 1180  
 QY 227 CTCCTCCAGATGGATGGATGTAGAAAAGTTGGAGAAGATGAGGGATCATGAA 280  
 DB 1181 GTACCTCCAGATGGATGGATGATATATTGGAGGAATCAAGGAGGAGAA 1234

Search completed: October 14, 2003, 14:17:03  
 Job time : 159.333 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:05:16 ; Search time 36.6667 Seconds  
(without alignments)  
3611.317 Million cell updates/sec

Title: US-09-854-124-6  
Perfect score: 300  
Sequence: 1 aattccgtgtgtgtgtga.....cgaaaggagaacttttac 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	38	12.7	1335	4	US-09-252-991A-5087
C 2	38	12.7	1389	4	US-09-252-991A-5133
C 3	34.8	11.6	1018	1	US-08-444-083-6
C 4	34.8	11.6	1018	1	US-08-286-304-6
C 5	34.8	11.6	1018	1	US-08-442-745-6
C 6	34.8	11.6	1018	1	US-08-443-129-6
C 7	34.8	11.6	1018	1	US-08-443-952-6
C 8	34.8	11.6	1018	1	US-08-443-130-6
C 9	34.8	11.6	1018	3	US-08-898-911-6
C 10	34.8	11.6	1018	5	PCT-US95-04467-6
C 11	34.8	11.6	1539	4	US-09-648-183-1
C 12	34.8	11.6	1539	4	US-09-648-183-2
C 13	33.6	11.2	429	4	US-09-252-991A-5055
C 14	33	11.0	711	4	US-09-252-991A-13138
C 15	33	11.0	783	4	US-09-252-991A-12804
C 16	32.6	10.9	480	4	US-09-073-009-135
C 17	32.6	10.9	1709	4	US-09-594-193-12
C 18	32.2	10.7	4403765	3	US-09-103-840A-2
C 19	32.2	10.7	4411529	3	US-09-103-840A-1
C 20	32	10.7	4403765	3	US-09-103-840A-2
C 21	32	10.7	4411529	3	US-09-103-840A-1
C 22	31.6	10.5	486	4	US-09-252-991A-6292
C 23	31.6	10.5	564	4	US-09-252-991A-6455
C 24	31.6	10.5	678	4	US-09-252-991A-2321
C 25	31.6	10.5	753	4	US-09-252-991A-2382
C 26	31.6	10.5	1241	1	US-08-471-033-39
C 27	31.6	10.5	1241	1	US-08-471-033-42

28	31.6	10.5	1241	2	US-08-471-044-39	Sequence 39, Appl
29	31.6	10.5	1241	2	US-08-471-044-42	Sequence 42, Appl
30	31.6	10.5	1241	2	US-08-463-483A-39	Sequence 39, Appl
31	31.6	10.5	1241	2	US-08-463-483A-42	Sequence 42, Appl
32	31.6	10.5	1241	2	US-08-471-046A-39	Sequence 39, Appl
33	31.6	10.5	1241	2	US-08-471-046A-42	Sequence 42, Appl
34	31.6	10.5	1241	2	US-08-470-566B-39	Sequence 39, Appl
35	31.6	10.5	1241	2	US-08-470-566B-42	Sequence 42, Appl
36	31.6	10.5	1241	2	US-08-469-334-39	Sequence 39, Appl
37	31.6	10.5	1241	2	US-08-469-334-42	Sequence 42, Appl
38	31.6	10.5	1241	3	US-09-300-529-39	Sequence 39, Appl
39	31.6	10.5	1241	3	US-09-300-529-42	Sequence 42, Appl
40	31.6	10.5	1358	1	US-08-471-033-45	Sequence 45, Appl
41	31.6	10.5	1358	2	US-08-471-044-45	Sequence 45, Appl
42	31.6	10.5	1358	2	US-08-463-483A-45	Sequence 45, Appl
43	31.6	10.5	1358	2	US-08-471-046A-45	Sequence 45, Appl
44	31.6	10.5	1358	2	US-08-470-566B-45	Sequence 45, Appl
45	31.6	10.5	1358	2	US-08-469-334-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1  
US-09-252-991A-5087/c  
; Sequence 5087, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 5087  
; LENGTH: 1335  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5087

Query Match	12.7%	Score 38;	DB 4;	Length 1335;
Best Local Similarity	51.8%	Pred. No. 0.061;		
Matches	86;	Conservative	0;	Mismatches 80; Indels 0; Gaps 0;
QY	10	GCTGTGCTGAGGCTCTGGCCTCGAGCTCGCGCCGCGCATGACGCTGCCGAGGTGCGAATT	69	
Db	421	GGTGGCATGGCGCTTCAGCTGCTGGAGATACCTGTACGGGGCGCTACTGGGGATG	362	
QY	70	CCTCGCGGAGAGAGCTGGTTACCATTTATCCCACTTCAGTCTGGACAGATCTACCT	129	
Db	361	CTTGGCGGAGTTGACGCGCTCGACCTTCGAGACCTTCTTACCGGCTGGAATTCG	302	
QY	130	CATCGGGGGGAGCTGGGGCTTTTAACTCTGTTTACCGTGGAA	175	
Db	301	CATCGGGCGCGCTGGCGGCTTCGACGCCGCGCCGCGGCGGAA	256	

RESULT 2  
US-09-252-991A-5133  
; Sequence 5133, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 5133  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-5133

Query Match 12.7%; Score 38; DB 4; Length 1389;  
Best Local Similarity 51.8%; Pred. No. 0.062;  
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 10 GCTGCTGGTGGGCTCTGCGCTGACGCTGCGCCGCCGATGACGCTGCGGAGGTGGAATT 69  
DB 1059 GGTGCGGATGGGCTTACGCTGGAGATACCTGTACGGCGCTACTGCGGATG 1118  
QY 70 CCTCGCCGAGAGGAGCTGGTTACCATTTATCCCAACTTACGCTGGGACAAGATCTACCT 129  
DB 1119 CTTGGCGAGTTCGACCGGCTGCACTTCGAGACCTGCTTCTACCAAGGCGCTGGAATTCGC 1178  
QY 130 CATCGGGGGGACCTGCGGCGCTTTTAAACCTGTGTTTACCGGTGGAA 175  
DB 1179 CATCGGCGCGGCTTGGCGGCTTCGACGCGCGGCCGAGGCGGAA 1224

RESULT 3  
US-08-444-083-6  
; Sequence 6, Application US/08444083  
; Patent No. 5571675  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Joffre  
; APPLICANT: Chien, Kenneth  
; APPLICANT: King, Kathleen  
; APPLICANT: Pennica, Diane  
; APPLICANT: Wood, William  
; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,083  
; FILING DATE: 17-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/233609  
; FILING DATE: 25-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/286304  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: 894PID5  
; TELEPHONE: 415/225-8674  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1018 bases

; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-444-083-6

Query Match 11.6%; Score 34.8; DB 1; Length 1018;  
Best Local Similarity 50.6%; Pred. No. 0.49;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 21 GGCTCTGGCTGACGCTGCGCGCCGATGAGGAGCTGCGGAGGTGCAATTCCTCGCCGAGA 80  
DB 484 GGGCGGAGCCCCCGCCGCCCTTACCGCGCTTCCGCCGCTTCCCGGCA 543  
QY 81 AGGAGCTGGTTACCATTTATCCCAACTTTCAGTCTGGACAAGATCTACCTCATCGGGGGG 140  
DB 544 AGGTGCTGGGCTCGCGCTTTGGCGGCTTACCGCGAGTGGCTGAGCCGCGCAGGAGGCG 603  
QY 141 ACCTGGGCGCTTTTAAACCTGTGTTTACCGGTGGAAGTGCCTCTG 186  
DB 604 ACCTGGGCGAGCTGCTGCGCGGGGCTCGGCTGAGCGCGCGGGG 649

RESULT 4  
US-08-286-304-6  
; Sequence 6, Application US/08286304  
; Patent No. 5571893  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Joffre  
; APPLICANT: Chien, Kenneth  
; APPLICANT: King, Kathleen  
; APPLICANT: Pennica, Diane  
; APPLICANT: Wood, William  
; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,304  
; FILING DATE: 05-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/233609  
; FILING DATE: 25-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 894P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1018 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-286-304-6

Query Match 11.6%; Score 34.8; DB 1; Length 1018;  
Best Local Similarity 50.6%; Pred. No. 0.49;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;



21	GGCTCTGGCTGCGACATCTGCGCGCGCATGAGACGCTCCGAGAGTGTGAATTCCTCTCCGCGAGA	80
QY		
484	GGCGAGACCCCGCCGCGCCACGCGCTACGCGCGCTCCGCCACCGGGGTTCTCCCGCGCA	543
Db		
81	AGGAGCTGGTTACCATATTACCCCAACTTCAGTCTGAGCAAGATCTACCTCATCGGGGGG	140
QY		
544	AGGTCTGGGCTCCGCGTTTTCGCGGCTCTACCGGAGTGGCTGAGCCGACCGAGGGCG	603
Db		
141	ACCTGGGGCCCTTTAAACCTCTGGTTTACCCGTGGAAGTGCCTCTGTG	186
QY		
604	ACCTGGGCGAGCTGCTGCGCGGGGGCTCGGCTGAGCGCGCGGGG	649
Db		

RESULT 5  
US-08-442-745-6  
; Sequence 6, Application US/08442745  
; Patent No. 5624806  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Joffre  
; APPLICANT: Chien, Kenneth  
; APPLICANT: King, Kathleen  
; APPLICANT: Pennica, Diane  
; APPLICANT: Wood, William  
; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor

	Query Match	11.6%	Score 34.8	DB 1	Length 1018
	Best Local Similarity	50.6%	Pred. No. 0.49		
	Matches	84	Conservative	0	Mismatches 82; Indels 0; Gaps 0;
QY	21	GCTCTGGCTCGAGCTCGCGCGCATGACGCTCGGAGTCGAATTCCTCGCCGAGA	80		
Db	484	GGCCGAGCCCCCGCGCCACCGCTTCAGCCGCTCCGCCACGGGGTCTTCCCCGCCA	543		
QY	81	AGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACACAGATCTACCTCATCGGGGGG	140		

```

Db      544  AGGTGCTGGGGCTCCGGCTTTGGCGCCCTTACCGCGAGTGCGGTAGCCGCCACCGAGGGCG 603
QY      141  ACCTGGGGCCCTTTTAAACCTCGTTTACCCGTGGGAAGTGCCCTGTG 186
          ||||| | | | | | | | | | | | | | | | | | | | |
Db      604  ACCTGGGCCAGCTCTGCCCGGGGGCTCGGCCTGAGCGCCGGGGG 649

RESULT 6
US-08-443-129-6
: Sequence 6, Application US/08443129
: Patent No. 5627073
: GENERAL INFORMATION:
: APPLICANT: Baker, Joffre
: APPLICANT: Chien, Kenneth
: APPLICANT: King, Kathleen
: APPLICANT: Pennica, Diane
: APPLICANT: Wood, William
: TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/443,129
: FILING DATE: 17-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/233609
: FILING DATE: 25-APR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/286
: FILING DATE: 05-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 304
: FILING DATE: (null)
: ATTORNEY/AGENT INFORMATION:
: NAME: Torchia, Timothy E.
: REGISTRATION NUMBER: 36,700
: REFERENCE/DOCKET NUMBER: 894PID2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-8674
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1018 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-443-129-6

Query Match      11.6%; Score 34.8; DB 1; Length 1018;
Best Local Similarity 50.6%; Pred. No. 0.49;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0

QY      21  GGCTCTGGCGTACCATTTATATCCCACTTCAGTCTGGACAGATCTACCTCATCGGGGGG 140
          || | | | | | | | | | | | | | | | | | | |
Db      484  GGGCCGAGCCCGCCGCCACCGCCTCAGCGCGCTCGCGCCACCGCGGGTCTTCCCGCCA 543
          || | | | | | | | | | | | | | | | | | | |

QY      81  AGGAGCTGGTTACCATTTATATCCCACTTCAGTCTGGACAGATCTACCTCATCGGGGGG 140
          || | | | | | | | | | | | | | | | | | | |
Db      544  AGGTGCTGGGGCTCCGGCTTTGGCGCCCTTACCGCGAGTGCGGTAGCCGCCACCGAGGGCG 603
QY      141  ACCTGGGGCCCTTTTAAACCTCGTTTACCCGTGGGAAGTGCCCTGTG 186

```

Db 604 ACCTGGGCGAGCTGCTGCCCGGGGCTCGGCTGAGCGCGGGG 649

## RESULT 7

US-08-443-952-6  
 ; Sequence 6, Application US/08443952  
 ; Patent No. 5679545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Joffre  
 ; APPLICANT: Chien, Kenneth  
 ; APPLICANT: King, Kathleen  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/443,952  
 ; FILING DATE: 17-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/233609  
 ; FILING DATE: 25-APR-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/286304  
 ; FILING DATE: 05-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Torchia, Timothy E.  
 ; REGISTRATION NUMBER: 36,700  
 ; REFERENCE/DOCKET NUMBER: 894PID4  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/225-8674  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1018 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-443-952-6

Query Match 11.6%; Score 34.8; DB 1; Length 1018;  
 Best Local Similarity 50.6%; Pred. No. 0.49;  
 Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
 QY 21 GGCTCTGGCTGACGCTGCCCGCCCATGGACGCTGCGAGGTGCAATTCTCGCCGAGA 80  
 Db 484 GGGCGGAGCCCCCGCCGACCCCTCAGCCGCTCGGCCACCGGGTCTTCCCGCCA 543  
 QY 81 AGGAGCTGGTTTACCAATTATCCCCAACTTACGCTCTGGACAAGACTACCTCATCGGGGG 140  
 Db 544 AGTGCTGGGGCTCCGGCTTTGCGGCTCTACCCGAGTGGCTGAGCGGACCGAGGGCG 603  
 QY 141 ACCTGGGCGCTTTTAACTGGTTTACCCGTTGGAAGTGCCCTGTG 186  
 Db 604 ACCTGGGCGAGCTGCTGCCCGGGGCTCGGCTGAGCGCGGGG 649

## RESULT 8

US-08-443-130-6

; Sequence 6, Application US/08443130  
 ; Patent No. 5723585  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Joffre  
 ; APPLICANT: Chien, Kenneth  
 ; APPLICANT: King, Kathleen  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/443,130  
 ; FILING DATE: 17-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/233609  
 ; FILING DATE: 25-APR-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/286304  
 ; FILING DATE: 05-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Torchia, Timothy E.  
 ; REGISTRATION NUMBER: 36,700  
 ; REFERENCE/DOCKET NUMBER: 894PID3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/225-8674  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1018 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-443-130-6

Query Match 11.6%; Score 34.8; DB 1; Length 1018;  
 Best Local Similarity 50.6%; Pred. No. 0.49;  
 Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
 QY 21 GGCTCTGGCTGACGCTGCCCGCCCATGGACGCTGCGAGGTGCAATTCTCGCCGAGA 80  
 Db 484 GGGCGGAGCCCCCGCCGACCCCTCAGCCGCTCGGCCACCGGGTCTTCCCGCCA 543  
 QY 81 AGGAGCTGGTTTACCAATTATCCCCAACTTACGCTCTGGACAAGACTACCTCATCGGGGG 140  
 Db 544 AGTGCTGGGGCTCCGGCTTTGCGGCTCTACCCGAGTGGCTGAGCGGACCGAGGGCG 603  
 QY 141 ACCTGGGCGCTTTTAACTGGTTTACCCGTTGGAAGTGCCCTGTG 186  
 Db 604 ACCTGGGCGAGCTGCTGCCCGGGGCTCGGCTGAGCGCGGGG 649

## RESULT 9

US-08-898-911-6  
 ; Sequence 6, Application US/08898911  
 ; Patent No. 6117650  
 ; GENERAL INFORMATION:  
 ; APPLICANT: King, Kathleen  
 ; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor



```

; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Human
US-09-648-183-1

```

[illegible]

RESULT 12  
US-09-648-183-2/c

```

: Sequence 2, Application US/09648183
: Patent No. 6472585
: GENERAL INFORMATION:
: APPLICANT: Botstein, David
: APPLICANT: Goddard, Audrey
: APPLICANT: Lawrence, David A.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret A.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Cardiotrophin-1 Compositions and Methods for the
: TREATMENT OF TUMOR
: FILE REFERENCE: P2533C1
: CURRENT APPLICATION NUMBER: US/09/648,183
: CURRENT FILING DATE: 2000-08-25
: PRIOR APPLICATION NUMBER: US 09/234,730
: PRIOR FILING DATE: 1999-01-21
: PRIOR APPLICATION NUMBER: US 09/033,114
: PRIOR FILING DATE: 1998-03-02
: PRIOR APPLICATION NUMBER: US 08/733,850
: PRIOR FILING DATE: 1996-10-18
: PRIOR APPLICATION NUMBER: US 08/443,129
: PRIOR FILING DATE: 1995-05-17
: PRIOR APPLICATION NUMBER: US 08/286,304
: PRIOR FILING DATE: 1994-08-05
: PRIOR APPLICATION NUMBER: US 08/233,609
: PRIOR FILING DATE: 1994-04-25
: PRIOR APPLICATION NUMBER: US 60/113,296
: PRIOR FILING DATE: 1998-12-22
: NUMBER OF SEQ ID NOS: 6
: SEQ ID NO 2
: LENGTH: 1539
: TYPE: DNA
: ORGANISM: Human
US-09-648-183-2

```

	Query Match Best Local Similarity Matches	11.6%; 50.6%; 84;	Score 34.8; Pred. No. 0.57; Conservative 0;	DB 4; Length 1539; Mismatches 82; Indels 0;	Gaps 0
QY	21	GGCTCTGGCTGACGTGCGCGCCGACGTGACGCGCTCGCGAGGTGCAATTCCTCGCCGAGA	80		
dDb	1056	GGGCGGAGCGCCCCCGCGCGCACGCGCTAGCGCGCTCGCGAGGTGCTTCCCGCGCA	997		
QY	81	AGGAGCTGGTTACCAATTATCCCAACTTCAGTCTGGACAGAGATCTACCTCATCGGGGGGGG	140		
dDb	996	AGGTGCTGGGGCTCCGCGTTTGGGGGCTCTACCGGAGTGGGTGAGCCGACACCGAGGGCG	937		

**QY** 141 ACCTGGGGCCCTTTTAAACCCCTGGTTTACCCCTGGAAGTGCCCTGTG 186  
||||||| | |||| | | |||||  
**Dd** 936 ACCTGGGCCAGCTGCTGCCCGGGGGCTCGGCCTGAGCGCCGGGGG 891

```

RESULT 13
US-09-252-991A-5055/c
; Sequence 5055, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5055
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5055

```

	Query Match	11.2%	Score 33.6;	DB 4;	Length 429;
	Best Local Similarity	51.3%;	Pred. No. 0.82;	Mismatches	78; Conservative
		0;	Gaps 0;	Indels 74;	Gaps 0;
QY	24	TCTGGCCTCGAGCTCGGCGCCCATGACGCTGCCGAGGTGGAATTCCTGCCCGAGAAGG	83		
Db	429				
QY	84	AGCTGGTTACCAATTATCCCCAACTCTAGTCTGGACAAGATCTACTCATCGGGGGGACC	143		
Db	369	ACGGGCTGCACCTTCGAGACCTTGCTTACACAGGCGCTTGGCAATTCGCCATCGGGCGCGGCC	310		
QY	144	TGGGGCCTTTTAACCCCTGGTTTTACCGGTGAA	175		
Db	309				
QY	309	TGCGCGGCTTCGACGCGGCGCGCGACGCCGAA	278		

```

RESULT 14
US-09-252-991A-13138
; Sequence 13138, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13138
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13138

```

Query Match	11.0%	Score 33;	DB 4;	Length 711;
Best Local Similarity	52.6%;	Pred. No. 1.5;		
Matches	72;	Conservative	0;	Mismatches 65; Indels 0; Gaps
Qy	11	CTGTGCTGAGGCTCTGGCCCTGCACGCTCCGCCGCCATGGACGCTGCCGAGGTCGCGAATTC	70	



**THIS PAGE BLANK (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 14:17:11 ; Search time 143.333 Seconds  
(without alignments)  
5431.000 Million cell updates/sec

Title: US-09-854-124-6

Perfect score: 300

Sequence: 1 aattccgtgtcgtcgtga.....cgaaagggaagaacttttac 300.

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	9	US-09-854-124-6
2	284.8	94.9	597	10	US-09-796-692-4407
3	284.8	94.9	597	14	US-10-040-862-4407
4	284.8	94.9	1210	9	US-09-925-301-214
5	126.2	42.1	484	11	US-09-918-995-19938
6	102.6	34.2	178	10	US-09-783-590-12142
7	63	21.0	382	14	US-09-796-692-9460
8	63	21.0	382	14	US-10-040-862-9460
9	36.4	12.1	3008	12	US-10-205-219-83
10	35.4	11.8	582	14	US-10-156-761-2179
11	35.4	11.8	9025608	14	US-10-156-761-1
12	34.8	11.6	1018	10	US-09-896-856-6
13	34.8	11.6	1018	14	US-10-107-931-6
14	34.8	11.6	1539	10	US-09-901-540-1
15	34.8	11.6	1539	10	US-09-901-540-2
16	34.8	11.6	1539	10	US-09-901-540-1
17	34.8	11.6	1539	10	US-09-901-257-1

c 17	34.8	11.6	1539	10	US-09-901-257-2	Sequence 2, Appli
18	34.8	11.6	1539	10	US-09-954-531-967	Sequence 967, App
19	34.8	11.6	3161	12	US-10-058-597-14	Sequence 14, Appl
c 20	34	11.3	501	14	US-10-156-761-7483	Sequence 7483, Ap
c 21	34	11.3	9025608	14	US-10-156-761-1	Sequence 1, Appli
22	33.6	11.2	32846	12	US-10-017-161-1557	Sequence 1557, Ap
23	33.4	11.1	1118	10	US-09-974-300-2728	Sequence 2728, Ap
c 24	33	11.0	2000	10	US-09-938-842A-5378	Sequence 135, App
c 25	32.6	10.9	480	9	US-09-073-009-135	Sequence 135, App
c 26	32.6	10.9	480	9	US-09-793-306-135	Sequence 6871, Ap
27	32.6	10.9	1695	14	US-10-156-761-6871	Sequence 12, Appl
28	32.6	10.9	1709	12	US-10-174-133-12	Sequence 3244, Ap
29	32.4	10.8	885	14	US-10-156-761-3244	Sequence 649, App
30	32.2	10.7	45191	15	US-10-080-170-2880	Sequence 2880, Ap
31	32	10.7	5721	14	US-10-156-761-15102	Sequence 15102, A
c 32	32	10.7	125746	14	US-10-156-761-631	Sequence 631, App
33	31.8	10.6	534	14	US-10-156-761-631	Sequence 41006, A
34	31.8	10.6	540	13	US-10-027-632-41006	Sequence 74601, A
35	31.8	10.6	546	13	US-10-027-632-74601	Sequence 299544,
36	31.8	10.6	546	13	US-10-027-632-299543	Sequence 299544,
37	31.8	10.6	546	13	US-10-027-632-299543	Sequence 85, Appl
38	31.8	10.6	978	12	US-09-971-392-85	Sequence 3799, Ap
39	31.8	10.6	1629	10	US-09-880-107-3799	Sequence 12, Appl
40	31.8	10.6	1629	10	US-09-880-107-3799	Sequence 33, Appl
41	31.8	10.6	1829	14	US-10-146-473-12	Sequence 25, Appl
42	31.8	10.6	10311	12	US-10-132-134-33	Sequence 1201, Ap
43	31.8	10.6	50543	12	US-10-132-134-25	Sequence 1917, Ap
44	31.6	10.5	379	10	US-09-880-107-1201	
45	31.6	10.5	930	14	US-10-156-761-1917	

#### ALIGNMENTS

#### RESULT 1

US-09-854-124-6  
Sequence 6, Application US/09854124  
Patent No. US20020076735A1

GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Escobedo, Jaime  
APPLICANT: Innis, Michael A.  
APPLICANT: Garcia, Pablo Dominguez  
APPLICANT: Sudduth-Klinger, Julie  
APPLICANT: Reinhard, Christoph  
APPLICANT: Giese, Klaus  
APPLICANT: Randazzo, Filippo  
APPLICANT: Kennedy, Giulia C.  
APPLICANT: Pot, David  
APPLICANT: Kassam, Altaf  
APPLICANT: Lamson, George  
APPLICANT: Drmanac, Radoje  
APPLICANT: Crkvenjakov, Radomir  
APPLICANT: Dickson, Mark  
APPLICANT: Drmanac, Snezana  
APPLICANT: Labat, Ivan  
APPLICANT: Leshkowitz, Dena  
APPLICANT: Kita, David  
APPLICANT: Garcia, Veronica  
APPLICANT: Jones, Lee William  
APPLICANT: Stache-Crain, Birgit  
TITLE OF INVENTION: Diagnostic and Therapeutic Methods Using  
TITLE OF INVENTION: Molecules Differentially Expressed in Cancer Cells  
FILE REFERENCE: 2300-1490  
CURRENT APPLICATION NUMBER: US/09/854,124  
CURRENT FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 09/400,947  
PRIOR FILING DATE: 1999-09-22  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 300  
TYPE: DNA

```
; ORGANISM: Homo sapiens
US-09-854-124-6

Query Match      100.0%; Score 300; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.6e-90;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATTCGGTGTCTCGGTGAGGCTCTGGCTGCAGCTCGCGGCCCATGGAGCTGCCGA 60
   |||||||
Db 1 AATTCGGTGTCTCGGTGAGGCTCTGGCTGCAGCTCGCGGCCCATGGAGCTGCCGA 60

QY 61 GGTCAAAATTCCTCGCGGAGAGGAGCTGTTACCAATTATCCCAACTTCAGTCTGACAA 120
   |||||||
Db 61 GGTCAAAATTCCTCGCGGAGAGGAGCTGTTACCAATTATCCCAACTTCAGTCTGACAA 120

QY 121 GATCTACCTCATCGCGGGGGGACCTGGGCGCTTTTAAACCTCGTGGTGAAGTGGC 180
   |||||||
Db 121 GATCTACCTCATCGCGGGGGGACCTGGGCGCTTTTAAACCTCGTGGTGAAGTGGC 180

QY 181 CTTGTGGCTGGGATTAACCTGAACAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTG 240
   |||||||
Db 181 CTTGTGGCTGGGATTAACCTGAACAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTG 240

QY 241 GATGATGTAGAAAAGTTGGAGAGATGAGGATCATGAACGAAAGGAAGAACTTTTAC 300
   |||||||
Db 241 GATGATGTAGAAAAGTTGGAGAGATGAGGATCATGAACGAAAGGAAGAACTTTTAC 300

RESULT 2
US-09-796-692-4407
; Sequence 4407, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4407
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (583)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (592)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-4407

Query Match      94.9%; Score 284.8; DB 10; Length 597;
Best Local Similarity 99.3%; Pred. No. 1.3e-84;
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 GTCGGTGAAGCTCTGGCTGCAGCTCGCGGCCCATGGAGCTGCCGAGTTCCT 72
   |||||||
Db 37 GACCGTGAAGCTCTGGCTGCAGCTCGCGGCCCATGGAGCTGCCGAGTTCCT 96

QY 73 CGCCGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 132
   |||||||
Db 97 CGCCGAGAAGGAGCTGGTTACCATTTATCCCAACTTCAGTCTGGACAGATCTACCTCAT 156

QY 133 CGGGGGGACCTGGGGCGCTTTTAAACCTGGTGTTCACCTGGAAATGCCCCCTGTGGCTGGC 192
   |||||||
Db 157 CGGGGGGACCTGGGGCGCTTTTAAACCTGGTGTTCACCTGGAAATGCCCCCTGTGGCTGGC 216

QY 193 GATTAACTGAAACAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGTAGA 252
   |||||||
Db 217 GATTAACTGAAACAAGACAGAAATGTGCGCTGCTCCCTCCAGAGTGGATGTAGA 276

QY 253 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAGAACTTTTAC 300
   |||||||
Db 277 AAAGTTGGAGAAGATGAGGATCATGAACGAAAGGAAGAACTTTTAC 324

RESULT 3
US-10-040-862-4407
; Sequence 4407, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040.862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999.
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
```





```

; SEQ ID NO 12142
; LENGTH: 178
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (49)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (61)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (66)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (143)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-12142

Query Match          34.2%; Score 102.6; DB 10; Length 178;
Best Local Similarity 80.2%; Pred. No. 4.2e-24;
Matches 142; Conservative 0; Mismatches 30; Indels 5; Gaps 2;

QY 47 ATGGACGTGCGAGGTGCAATTCCTCGCGAGAGGAGCTGGTTACCAATTATCCCCAAC 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 ATGAACGTGCGCAAGGTCAAAATCTCTCGCCGAAAGGAGCTGGTTACNATTATCCCCAAN 61

QY 107 TTCAGTCTGGCAAGATCTACTCATCGGGGGGGA--CCTGGGGCCTTTTAAACCTGGTT 164
   || || || || || || || || || || || || || || || || || || || || ||
Db 62 TTAATTTGGCAAAATCTACTATATGGGGGGGAACCTGGGGCCTTTTAAACCTGGTT 121

QY 165 TACCGGTGAA---GTGCCCCCTGTGGTGGCGATTAACTGAACAAAGACAGAGAAAT 218
   ||||| ||| || || || || || || || || || || || || || || || || || ||
Db 122 TACCGGTGGAGTGCCCTTGTGGTGGGAATTAACCTGAANAAGNCAGAAAT 178

RESULT 7
US-09-796-692-9460/c
; Sequence 9460, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
;
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9460
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (72)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (80)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (89)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (95)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (122)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (129)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (131)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (155)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (173)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (222)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (235)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (236)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (237)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (245)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (251)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (261)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (305)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (312)
; OTHER INFORMATION: n=A,T,C or G

```

```
; NAME/KEY: unsure
; LOCATION: (321)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (334)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (352)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (354)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (373)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (375)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-9460

Query Match      21.0%; Score 63; DB 10; Length 382;
Best Local Similarity 91.3%; Pred. No. 9e-11;
Matches 63; Conservative 0; Mismatches 6; Indels 0; Caps 0;

QY 232 TCCAGAGTGGATGATGAGAAAGTTGGAGAGATGAGGGATCATGACGAAAGGAAGA 291
Db 382 TCCAGAGNATGGATGATGAGAAAGTTGNANAAGATGAGGGATCATGACGAAAGGAAGA 323
QY 292 AACTTTTAC 300
Db 322 AACTTTTAC 314

RESULT 8
US-10-040-862-9460/c
; Sequence 9460, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NAME/KEY: unsure
; LOCATION: (321)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (334)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (352)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (354)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (373)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (375)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-9460
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9460
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (72)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (80)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (89)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (95)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (122)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (129)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (131)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (155)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (173)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (222)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (235)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (236)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (237)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (245)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (251)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
```

```

: LOCATION: (261)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (305)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (312)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (321)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (334)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (352)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (354)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (373)
: OTHER INFORMATION: n=A,T,C or G
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (375)
: OTHER INFORMATION: n=A,T,C or G
US-10-040-862-9460

```

	Query Match	21.0%;	Score 63;	DB 14;	Length 382;
	Best Local Similarity	91.3%;	Pred. No. 9e-11;		
	Matches 63;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	232	TCGACACTGGATGGATGTAGAAAAGTTGGAGAACATGAGGATCATGAAACGAAAGGAAGA	291		
Db	382	TCCAGAGNAGTGGATGTAGAAAAGTTGANAAGATGAGGATCATGANCAGAAAGGAAGA	323		
QY	292	AACTTTTAC	300		
Db	322	ANCTTTTAC	314		

```

RESULT 9
US-10-205-219-83/c
; Sequence 83, Application US/10205219
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 3008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: N-G, N-G dimethylarginine dimethylaminohydrolase

```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2411, 2412
; OTHER INFORMATION: n is a or g or c or t
US-10-205-219-83

Query Match      12.1%; Score 36.4; DB 12; Length 3008;
Best Local Similarity 46.2%; Pred. No. 0.15;
Matches 121; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

Qy      17 GTGAGGCTCGGCCTGTGAGCTCCGGCGGCATGACGCTGCCGAGGTTCGAATTCTCTCGCC 76
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      740 GTCAACTCTCTCTCTGGGCTAGCGGCCCGGGGGGTGATGAGGGCGGTCTCTTCGCA 681

Qy      77 GAGAAGAGTGTTTACCATTATCCCAACTTTCAGTCTGGACAAGATPACCTCATCGGG 136
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      680 CACCACGGGCACGCTCCTCCACGAACACGAGTCAGGCAGGCTCTCTCGCGCGGACGTCG 621

Qy      137 GGGGACCTGGGGCCTTTTAACCTGTGTTTACCCTGGAAAGTGCCCTGTGGCTGGCGATT 196
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      620 CACCACCTGGAGCCCCAGCTTGTGTGCCACGACGCCACGCTAGAGCTGTGTGTGGCGCTC 561

Qy      197 AACCTGAAAAAAGACAGAAAATGTCGCTGCTCCCTCCAGAGTGGAGTGGATGTAGAAAAG 256
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      560 AGCGGAGCGAATCACCTCTTCGCCCTGGGAGCGCCTCAGCGCTGCGGCACAGGA 501

Qy      257 TTGGAGAAGATGAGGGATCATG 278

Db      500 CTCGGGCGGAGCCGCCACACG 479
```

RESULT 10  
US-10-156-761-2179  
; Sequence 2179, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 2179  
; LENGTH: 582  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(582)  
US-10-156-761-2179

Query Match	11.8%;	Score 35.4;	DB 14;	Length 582;
Best Local Similarity	63.5%;	Pred. No. 0.17;		
Matches	54;	Conservative 0;	Mismatches 31;	Indels 0; Gaps 0;
QY	40	CGCGGCCCATGGAGCTGTCGAGGTGCGAATTC	TCCGCCGAGAAGGAGCTGGTTACCAATTAT	99
Dd	159	CGCGCGCCTGGACGCTCGACGACTCGACGGATCGG	GAGAAGGAGCCACCACCGTGGC	218
QY	100	CCCCAACTTCAGTCTGGACAAGATC		124
Dd	219	CCTCGACTTCGGCCGGATCACGATC		243

```
RESULT 11
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match      11.8%; Score 35.4; DB 14; Length 9025608;
Best Local Similarity 63.5%; Pred. No. 7.6;
Matches 54; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 40 CGCGCCATGGAGCTGCGGAGTGAATTCCTCGCGAGAGGAGCTGGTTACCATTTAT 99
Db 2665457 CGCGCGCCCTGGAGCTGAGAGCTCGAGCGATCGCGGAGAGAGCCACCACCGTGGC 2665516

QY 100 CCCCACCTTCAGTCTGGACAAGATC 124
Db 2665517 CCGTGACTTCGCGGATCAGCATC 2665541

RESULT 12
US-09-896-856-6
; Sequence 6, Application US/09896856
; Patent No. US20020137189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Joffre
; Chien, Kenneth
; King, Kathleen
; Pennica, Diane
; Wood, William
; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/896,856
; FILING DATE: 29-Jun-2001
; CLASSIFICATION: <Unknown>
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/733,850
FILING DATE: 18-OCT-1996
APPLICATION NUMBER: US 08/471,112
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: 08/233,609
FILING DATE: 25-APR-1994
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: 08/443129
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P0894PID2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-2066
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1018 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-896-856-6

Query Match      11.6%; Score 34.8; DB 10; Length 1018;
Best Local Similarity 50.6%; Pred. No. 0.33;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 21 GGCTCTGGCTGCAGCTCGCGCCATGGAGCTGCCGAGGTGCGAATTCCTCGCGGAGA 80
Db 484 GGGCGAGCCCCCGCGCCACCGCTCAGCGCGCTCCGCCACCGGGGTCTTCCCGCCA 543

QY 81 AGAGCTGTTTACCATTTATCCCAACTTCAGTCTGGACAAGATCTACCTCATCGGGGGG 140
Db 544 AGGTGTGGGGTCTCGCGCTTTGGCGCTCTACCGCGAGTGGCTGAGCCCGCAGCGGCG 603

QY 141 ACCTGGGGCTTTTAAACCTGGTTTACCCTGTTTACCGTGAAGTGCCTGTG 186
Db 604 ACCTGGGCGAGCTGCTGCCCGGGGCTCGCGCTGAGCCCGCGGGG 649

RESULT 13
US-10-107-931-6
; Sequence 6, Application US/10107931
; Publication No. US20030054550A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Joffre
; Chien, Kenneth
; King, Kathleen
; Pennica, Diane
; Wood, William
; TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/107,931
; FILING DATE: 26-Mar-2002
; CLASSIFICATION: <Unknown>
```

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/364,187  
; FILING DATE: 30-Jul-1999  
; APPLICATION NUMBER: 08/286,304  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 894P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1018 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-107-931-6

Query Match 11.6%; Score 34.8; DB 14; Length 1018;  
Best Local Similarity 50.6%; Pred. No. 0.33;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 21 GGCTCTGGCCTGCAGCTCGCGCCGCAATGACGCTGCGAGGTGGAATTCCTCGCCGAGA 80  
DB 484 GGGCGAGCGCCCGCCGACCGCTCAGCGCTCGCCACCGGGTCTTCCCGCCA 543  
QY 81 AGGAGCTGGTTACCAATATCCCAATCTCAGTCTGGACAAGATCTACCTCATCGGGGGG 140  
DB 544 AGGTGCTGGGCTCGCGCTTTGCGGCTCTACCGGAGTGTGAGCGCCGACCGAGGCG 603  
QY 141 ACCTGGGCGCTTTAAACCTGTTTACCGTGGAGTGCCTCTGTG 186  
DB 604 ACCTGGGCGAGCTGCTGCGCGGGGCTCGGCTGAGCGCGCGGGG 649

RESULT 14  
US-09-901-540-1  
; Sequence 1, Application US/09901540  
; Patent No. US20020102622A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret A.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Cardiotrophin-1 Compositions and Methods for the  
; FILE REFERENCE: P253D2  
; CURRENT APPLICATION NUMBER: US/09/901,540  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/648,258  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: US 09/033,114  
; PRIOR FILING DATE: 1998-03-02  
; PRIOR APPLICATION NUMBER: US 08/733,850  
; PRIOR FILING DATE: 1996-10-18  
; PRIOR APPLICATION NUMBER: US 08/443,129  
; PRIOR FILING DATE: 1995-05-17  
; PRIOR APPLICATION NUMBER: US 08/286,304  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: US 08/233,609  
; PRIOR FILING DATE: 1994-04-25  
; PRIOR APPLICATION NUMBER: US 60/113,296  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 6  
; SEQ ID NO 1  
; LENGTH: 1539  
; TYPE: DNA  
; ORGANISM: Human  
US-09-901-540-2

Query Match 11.6%; Score 34.8; DB 10; Length 1539;  
Best Local Similarity 50.6%; Pred. No. 0.39;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 21 GGCTCTGGCCTGCAGCTCGCGCCGCAATGACGCTGCGAGGTGGAATTCCTCGCCGAGA 80  
DB 1056 GGGCGAGCGCCCGCCGACCGCTCAGCGCTCGCCACCGGGTCTTCCCGCCA 997  
QY 81 AGGAGCTGGTTACCAATATCCCAATCTCAGTCTGGACAAGATCTACCTCATCGGGGGG 140  
DB 996 AGGTGCTGGGCTCGCGCTTTGCGGCTCTACCGGAGTGTGAGCGCCGACCGAGGCG 937  
QY 141 ACCTGGGCGCTTTAAACCTGTTTACCGTGGAGTGCCTCTGTG 186  
DB 936 ACCTGGGCGAGCTGCTGCGCGGGGCTCGGCTGAGCGCGCGGGG 891

; ORGANISM: Human  
US-09-901-540-1  
Query Match 11.6%; Score 34.8; DB 10; Length 1539;  
Best Local Similarity 50.6%; Pred. No. 0.39;  
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 21 GGCTCTGGCCTGCAGCTCGCGCCGCAATGACGCTGCGAGGTGGAATTCCTCGCCGAGA 80  
DB 484 GGGCGAGCGCCCGCCGACCGCTCAGCGCTCGCCACCGGGTCTTCCCGCCA 543  
QY 81 AGGAGCTGGTTACCAATATCCCAATCTCAGTCTGGACAAGATCTACCTCATCGGGGGG 140  
DB 544 AGGTGCTGGGCTCGCGCTTTGCGGCTCTACCGGAGTGTGAGCGCCGACCGAGGCG 603  
QY 141 ACCTGGGCGCTTTAAACCTGTTTACCGTGGAGTGCCTCTGTG 186  
DB 604 ACCTGGGCGAGCTGCTGCGCGGGGCTCGGCTGAGCGCGCGGGG 649

RESULT 15  
US-09-901-540-2/C  
; Sequence 2, Application US/09901540  
; Patent No. US20020102622A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret A.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Cardiotrophin-1 Compositions and Methods for the  
; FILE REFERENCE: P253D2  
; CURRENT APPLICATION NUMBER: US/09/901,540  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/648,258  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: US 09/033,114  
; PRIOR FILING DATE: 1998-03-02  
; PRIOR APPLICATION NUMBER: US 08/733,850  
; PRIOR FILING DATE: 1996-10-18  
; PRIOR APPLICATION NUMBER: US 08/443,129  
; PRIOR FILING DATE: 1995-05-17  
; PRIOR APPLICATION NUMBER: US 08/286,304  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: US 08/233,609  
; PRIOR FILING DATE: 1994-04-25  
; PRIOR APPLICATION NUMBER: US 60/113,296  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 6  
; SEQ ID NO 2  
; LENGTH: 1539  
; TYPE: DNA  
; ORGANISM: Human  
US-09-901-540-2

Search completed: October 14, 2003, 17:17:44  
Job time : 154.333 secs

---

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 13:00:56 ; Search time 1477 Seconds  
(without alignments)  
8309.335 Million cell updates/sec

Title: US-09-854-124-5  
Perfect score: 300  
Sequence: 1 acgaataccgacccctggtc.....tgaggtaactcgtggtctgg 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues 5777422  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.ma.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg\_hum.\*

31: em.htg\_inv.\*

32: em.htg\_other.\*

33: em.htg\_mus.\*

34: em.htg\_pln.\*

35: em.htg\_rod.\*

36: em.htg\_nam.\*

37: em.htg\_vrt.\*

38: em\_sy.\*

39: em.htgo\_hum.\*

40: em.htgo\_mus.\*

41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	300	100.0	858	6	BD156103	BD156103 Primer fo
2	300	100.0	858	9	AK001275	AK001275 Homo sapi
3	300	100.0	1174	9	AF151880	AF151880 Homo sapi
4	300	100.0	1180	9	BC010164	BC010164 Homo sapi
5	300	100.0	1188	9	AF201939	AF201939 Homo sapi
6	298.4	99.5	1203	9	BC003186	BC003186 Homo sapi
7	292.4	97.5	667	6	BD147616	BD147616 Primer fo
C 8	258.4	86.1	116792	9	HS1100H13	AL035419 Human DNA
C 9	201	67.0	160356	9	AC009068	AC009068 Homo sapi
C 10	201	67.0	185321	9	AC123908	AC123908 Homo sapi
C 11	201	67.0	245210	2	AC137771	AC137771 Homo sapi
C 12	199.4	66.5	3097	9	AK091519	AK091519 Homo sapi
C 13	140	46.7	454	6	AX336730	AX336730 Sequence
C 14	125.8	41.9	933	5	AB097169	AB097169 Xenopus l
C 15	111.8	37.3	207307	2	AC117085	AC117085 Rattus no
C 16	104.4	34.8	188292	2	AC103360	AC103360 Mus muscu
C 17	66	22.0	362	6	AX396467	AX396467 Sequence
C 18	53.8	17.9	825	3	AY071234	AY071234 Drosophil
C 19	53.8	17.9	72722	3	AC004373	AC004373 Drosophil
C 20	53.8	17.9	132666	3	AC005270	AC005270 Drosophil
C 21	53.8	17.9	182897	3	AC092232	AC092232 Drosophil
C 22	53.8	17.9	218794	2	AC018064	AC018064 Drosophil
C 23	53.8	17.9	292393	3	AE003576	AE003576 Drosophil
C 24	44.6	14.9	165158	2	AC141687	AC141687 Apis mell
C 25	41.6	13.9	213	9	AF534827	AF534827 Homo sapi
C 26	41.6	13.9	92814	9	AL391903	AL391903 Human DNA
C 27	41.6	13.9	210842	2	AC140748	AC140748 Rattus no
C 28	41.6	13.9	219443	2	AC023320	AC023320 Homo sapi
C 29	41.6	13.9	246542	2	AC125601	AC125601 Rattus no
C 30	39.6	13.2	734	6	BD150236	BD150236 Primer fo
C 31	39.6	13.2	1150	9	AF534824	AF534824 Homo sapi
C 32	39.6	13.2	3071	6	BD158620	BD158620 Primer fo
C 33	39.6	13.2	3071	9	AK023755	AK023755 Homo sapi
C 34	39.6	13.2	91511	9	AL133404	AL133404 Human DNA
C 35	39.6	13.2	92814	9	AL391903	AL391903 Human DNA
C 36	36.2	12.1	162641	2	AC091406	AC091406 Rattus no
C 37	36	12.0	654	6	AX386100	AX386100 Sequence
C 38	36	12.0	49995	2	AC011335	AC011335 Homo sapi
C 39	36	12.0	77139	2	AC008466	AC008466 Homo sapi
C 40	36	12.0	120849	2	AC087148	AC087148 Mus muscu
C 41	36	12.0	128928	9	AC068658	AC068658 Homo sapi
C 42	36	12.0	157208	9	AC113409	AC113409 Homo sapi
C 43	36	12.0	203418	2	AC063958	AC063958 Homo sapi
C 44	35.6	11.9	1129	8	AF499724	AF499724 Thellungi
C 45	35.6	11.9	215707	2	AC099096	AC099096 Rattus no

ALIGNMENTS

RESULT 1  
BD156103  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

BD156103  
Primer for synthesizing full-length cDNA and use thereof.  
BD156103  
BD156103.1 GI:27861861  
JP 2002191363-A/10946.  
Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 858)

Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

858 bp DNA linear RNT 17-JAN-2003

JOURNAL Patent: JP 2002191363-A 10946 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002191363-A/10946  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA, TAKAO ISOgai, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 PI KEIICHI NAGAI, TETSUJI OTSUKI  
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC  
 10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 PC Primer for synthesizing full-length cDNA and use thereof FH key  
 Location/Qualifiers  
 FT CDS (37) ..(591).  
 Location/Qualifiers  
 SOURCE 1..858  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 227 a 214 c 222 g 195 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 6; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-69;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGTCAGGATATGTGGGACATCGTATAGCCAACTCCGAGTGT 60  
 DB 368 ACGAAATCCGACCCCTGTCAGGATATGTGGGACATCGTATAGCCAACTCCGAGTGT 427  
 QY 61 CTGCTGACAGCTTTGTGAGCAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 120  
 DB 428 CTGCTGACAGCTTTGTGAGCAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 487  
 QY 121 TGGAGATCAACACACAGCGGGACTTCTCACAAGCGCTCAACCCATGTACAACTCC 180  
 DB 488 TGGAGATCAACACACAGCGGGACTTCTCACAAGCGCTCAACCCATGTACAACTCC 547  
 QY 181 GCAGAACTCCAGCCCTGTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 548 GCAGAACTCCAGCCCTGTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 607  
 QY 241 GCAGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCTGTTCTGG 300  
 DB 608 GCAGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCTGTTCTGG 667  
 RESULT 2  
 AK001275 858 bp mRNA linear PRI 01-AUG-2002  
 LOCUS Homo sapiens cDNA FLJ10413 fis, clone NT2RP1000063.  
 DEFINITION AK001275  
 ACCESSION  
 VERSION AK001275.1 GI:7022428  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
 Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
 Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
 Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
 Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 REFERENCE 2 (bases 1 to 858)  
 AUTHORS Isogai, T. and Otsuki, T.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Direct Submission  
 Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction, 5'- & 3'-end one pass sequencing and clone selection;  
 Helix Research Institute (supported by Japan Key Technology Center  
 etc.) and Department of Virology, Institute of Medical Science,  
 University of Tokyo.  
 FEATURES Location/Qualifiers  
 source 1..858  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP1000063"  
 /cell\_line="NT2"  
 /cell\_type="teratocarcinoma"  
 /clone\_lib="NT2RP1"  
 /note="cloning vector: pUC19FL3-mRNA from NT2 neuronal  
 precursor cells after 48-hours retinoic acid (RA)  
 induction."  
 CDS 37..594  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAA91595.1"  
 /db\_xref="GI:7022429"  
 /translation="MDAAVEFLAEKELVITIPNFDLKIYLGDDLPNPGLPVEV  
 PLMLAIKQKRCRLPPEDMDVEKLEMRHERKEETFPMPSPYIMETLKLLNH  
 ASDNIPKAEIRTLVDMWDTRIAKRVSDSEFVRQOEHAHKLNLTIMEINTSGTEL  
 TQALNNHYKLRTNLQPLESTQSDP"  
 BASE COUNT 227 a 214 c 222 g 195 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 9; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-69;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGTCAGGATATGTGGGACATCGTATAGCCAACTCCGAGTGT 60  
 DB 368 ACGAAATCCGACCCCTGTCAGGATATGTGGGACATCGTATAGCCAACTCCGAGTGT 427  
 QY 61 CTGCTGACAGCTTTGTGAGCAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 120  
 DB 428 CTGCTGACAGCTTTGTGAGCAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 487  
 QY 121 TGGAGATCAACACACAGCGGGACTTCTCACAAGCGCTCAACCCATGTACAACTCC 180  
 DB 488 TGGAGATCAACACACAGCGGGACTTCTCACAAGCGCTCAACCCATGTACAACTCC 547  
 QY 181 GCAGAACTCCAGCCCTGTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 548 GCAGAACTCCAGCCCTGTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 607  
 QY 241 GCAGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCTGTTCTGG 300  
 DB 608 GCAGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCTGTTCTGG 667  
 RESULT 3  
 AF151880 1174 bp mRNA linear PRI 18-MAY-2000  
 LOCUS Homo sapiens CGI-122 protein mRNA, complete cds.  
 DEFINITION AF151880  
 ACCESSION AF151880  
 VERSION AF151880.1 GI:4929712  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1174)  
 AUTHORS Lai, C.H., Chou, C.Y., Ch'ang, L.Y., Liu, C.S. and Lin, W.

**TITLE** Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics

**JOURNAL MEDLINE** Genome Res. 10 (5), 703-713 (2000)

**PUBMED** 2072150

**REFERENCE** 10810093

**AUTHORS** 2 (bases 1 to 1174)

**TITLE** Direct Submission

**JOURNAL** Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan

**FEATURES** Location/Qualifiers

1..1174

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="16"

77..634

/codon\_start=1

/product="CGI-122 protein"

/protein\_id="AD34117.1"

/db\_xref="GI:4929713"

/translation="MDAAEVEFLAEKELVTIIPNFSLDKIYLGDLGPFNPLPVEV PLWLAINKQKOKRLLPPWMDVLEKMRDHERKEETFTPMPSPYMYELTKLLLNH ASDNIPKADEIRTLVKMDWDRIAKLRSADSFVRQEAHAKLDNLTLMEINTSGTFL TQALNHYKLRITNLPLESTQSDP"

**BASE COUNT** 300 a 299 c 307 g 268 t

**ORIGIN**

Query Match 100.0%; Score 300; DB 9; Length 1174;

Best Local Similarity 100.0%; Pred. No. 4.5e-69;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAATCCGGACCTGTCAAGGATATGTGGACACTCGTATAGCCAAATCCGAGTGT 60

DB 408 ACGAATCCGGACCTGTCAAGGATATGTGGACACTCGTATAGCCAAATCCGAGTGT 467

QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATTAATCTGACCTTGA 120

DB 468 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATTAATCTGACCTTGA 527

QY 121 TGGAGATCAACACACAGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAATCC 180

DB 528 TGGAGATCAACACACAGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAATCC 587

QY 181 GCACGAATCCAGCCTCTGGAGAGTACTCACTCTCAGAGCTTCAGAGAAAGGCGTGT 240

DB 588 GCACGAATCCAGCCTCTGGAGAGTACTCACTCTCAGAGCTTCAGAGAAAGGCGTGT 647

QY 241 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTGG 300

DB 648 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTGG 707

**RESULT 4**

**BC010164**

**LOCUS** BC010164 1180 bp mRNA linear PRI 12-JUL-2001

**DEFINITION** Homo sapiens, HSPC037 protein, clone MGC:19836 IMAGE:4098007, mRNA, complete cds.

**ACCESSION** BC010164

**VERSION** BC010164.1 GI:14603431

**KEYWORDS** MGC.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1180)

**REFERENCE** Strausberg, R.

**AUTHORS** Direct Submission

**TITLE** Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**JOURNAL**

**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>

## COMMENT

Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 27 Row: 1 Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.

## FEATURES

Location/Qualifiers

1..1180

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="LocusID:51659"

/db\_xref="taxon:9606"

/clone="MGC:19836 IMAGE:4098007"

/tissue\_type="Brain, neuroblastoma"

/clone\_lib="NIH\_MGC\_19"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

68..625

/codon\_start=1

/product="HSPC037 protein"

/protein\_id="AAH0164.1"

/db\_xref="GI:14603432"

/translation="MDAAEVEFLAEKELVTIIPNFSLDKIYLGDLGPFNPLPVEV PLWLAINKQKOKRLLPPWMDVLEKMRDHERKEETFTPMPSPYMYELTKLLLNH ASDNIPKADEIRTLVKMDWDRIAKLRSADSFVRQEAHAKLDNLTLMEINTSGTFL TQALNHYKLRITNLPLESTQSDP"

## CDS

**BASE COUNT** 316 a 292 c 307 g 265 t

**ORIGIN**

Query Match 100.0%; Score 300; DB 9; Length 1180;

Best Local Similarity 100.0%; Pred. No. 4.5e-69;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAATCCGGACCTGTCAAGGATATGTGGACACTCGTATAGCCAAATCCGAGTGT 60

DB 399 ACGAATCCGGACCTGTCAAGGATATGTGGACACTCGTATAGCCAAATCCGAGTGT 458

QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATTAATCTGACCTTGA 120

DB 459 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATTAATCTGACCTTGA 518

QY 121 TGGAGATCAACACACAGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAATCC 180

DB 519 TGGAGATCAACACACAGCGGACTTTCCTCACACAGCGCTCAACACATGTACAAATCC 578

QY 181 GCACCAATCCAGCCTCTGGAGAGTACTCACTCTCAGGAGCTTCTAGAGAAAGGCGTGT 240

DB 579 GCACCAATCCAGCCTCTGGAGAGTACTCACTCTCAGGAGCTTCTAGAGAAAGGCGTGT 638

QY 241 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTGG 300

DB 639 GCAGCGGCTTCTCGGGGATGTGAGCGCTCAGGAGCTGATGAGTACTCGTGGTCTGG 698

## RESULT 5

AF201939  
LOCUS AF201939 Homo sapiens DC5 (DC5) mRNA, complete cds. linear PRI 20-JUL-2000  
DEFINITION AF201939 Homo sapiens DC5 (DC5) mRNA, complete cds.  
ACCESSION AF201939  
VERSION AF201939.1 GI:9295181  
KEYWORDS FLI\_CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1188)  
AUTHORS Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.  
TITLE Novel genes expressed in human dentritic cell  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1188)  
AUTHORS Li,Y., Peng,Y., Li,N., Gu,W., Han,Z., Fu,G. and Chen,Z.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
FEATURES  
Location/Qualifiers  
1..1188  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_type="dentritic cell"  
1..1188  
/gene="DC5"  
77..634  
/gene="DC5"  
/codon\_start=1  
/product="DC5"  
/protein\_id="AAF86875.1"  
/db\_xref="GI:9295181"  
/translation="MDAAVEFLAEKELVTIIPNFSLDKIYLGDLGPNPGLPVEV  
PLWAINLKQKORCLLPPEMDVEKLEKMRDHERKEETTPMPSPYMBELTKLLLNH  
ASDNIPKADEIRTLVDMWDTRIAKLRVSDSFVRQEAHAKLDNLTLMEINTSGTFL  
TQALNHMYKLTNLPLESTQSQDF"  
BASE COUNT 314 a 299 c 307 g 268 t  
ORIGIN  
Query Match 100.0%; Score 300; DB 9; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 4.5e-69;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACGAAATCCGGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
Db AAAAAAAAAACCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 67  
Qy 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 120  
Db CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 127  
Qy 121 TGGAGATCAACACACAGCGGACCTTTCCTCACACAAGCGCTCAACCATGTACAAACTCC 180  
Db TGGAGATCAACACACAGCGGACCTTTCCTCACACAAGCGCTCAACCATGTACAAACTCC 187  
Qy 181 GCAGACACCTCCAGCCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
Db GCAGACACCTCCAGCCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 247  
Qy 241 GCAGGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTCTGG 300  
Db GCAGGCGGCTTGTCTGGGGATGTGAGCGCTCAGGACGTGATGAGTACTCGTGGTCTGG 307  
RESULT 6  
LOCUS BC003186 1203 bp mRNA linear PRI 12-JUL-2001  
DEFINITION BC003186 Homo sapiens, HSPC037 protein, clone MGC:3503268, mRNA, complete cds.  
ACCESSION BC003186  
VERSION BC003186.1 GI:13112024

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT

MGC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1203)  
Straussberg,R.  
Direct Submission  
Submitted (13-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Seedei, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 6 Row: d Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4929712.

FEATURES  
source

Location/Qualifiers  
1..1203  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:673 IMAGE:3503268"  
/tissue\_type="Brain, neuroblastoma"  
/clone\_lib="NIH\_MGC\_19"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
85..642  
/codon\_start=1  
/product="HSPC037 protein"  
/protein\_id="AAH03186.1"  
/db\_xref="GI:13112025"  
/translation="MDAAVEFLAEKELVTIIPNFSLDKIYLGDLGPNPGLPVEV  
PLWAINLKQKORCLLPPEMDVEKLEKMRDHERKEETTPMPSPYMBELTKLLLNH  
ASDNIPKADEIRTLVDMWDTRIAKLRVSDSFVRQEAHAKLDNLTLMEINTSGTFL  
TQALNHMYKLTNLPLESTQSQDF"  
BASE COUNT 326 a 300 c 309 g 268 t  
ORIGIN

Query Match 99.5%; Score 298.4; DB 9; Length 1203;  
Best Local Similarity 99.7%; Pred. No. 1.2e-68;  
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ACGAAATCCGGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
Db AAAAAAAAAACCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 67  
Qy 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 120  
Db CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCCAAGCTGGATAACTTGACCTTGA 127  
Qy 121 TGGAGATCAACACACAGCGGACCTTTCCTCACACAAGCGCTCAACCATGTACAAACTCC 180  
Db TGGAGATCAACACACAGCGGACCTTTCCTCACACAAGCGCTCAACCATGTACAAACTCC 187

QY 181 GCACGAACCTCCAGCCTCTGGAGATCTACTCAGTCTCAGGACTCTTAGAGAAAGCGCTGT 240  
|||||  
Db 596 GCACAAACCTCCAGCCTCTGGAGATCTACTCAGTCTCAGGACTCTTAGAGAAAGCGCTGT 655  
QY 241 GCAGCGGCTCTCTGGGGATCTGAGCGCTCAGGAGTATGATCAGTACTCTGGTCTTGG 300  
|||||  
Db 656 GCAGCGGCTCTCTGGGGATCTGAGCGCTCAGGAGTATGATCAGTACTCTGGTCTTGG 715  
RESULT 7  
BD147616  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002191363-A/2459  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT source 1..667  
FT Location/Qualifiers  
1..667  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 171 a 176 c 177 g 137 t 6 others  
ORIGIN  
Query Match 97.5%; Score 292.4; DB 6; Length 667;  
Best Local Similarity 97.7%; Pred. No. 5.2e-67;  
Matches 293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 ACGAATCCGACCTCTGTCAGGATATGTGGGACATCTGTATAGCAAACTCCGAGTGT 60  
|||||  
Db 367 ACGAATCCGACCTCTGTCAGGATATGTGGGACATCTGTATAGCAAACTCCGAGTGT 426  
QY 61 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACTTGA 120  
|||||  
Db 427 CTGCTGACAGCTTTGTGACAGCAGGAGGACATGCCAAGCTGGATACCTTGACTTGA 486  
QY 121 TGGAGATCAACACCGCGGACTTTCTCCACACAAGCGCTCAACCATCTGTACAACTCC 180  
|||||  
Db 487 TGGAGATCAACACCGCGGACTTTCTCCACACAAGCGCTCAACCATCTGTACAACTCC 546  
QY 181 GCACGAACCTCCAGCCTCTGGAGATCTACTCAGTCTCAGGACTCTTAGAGAAAGCGCTGT 240  
|||||  
Db 547 GCACGAACCTCCAGCCTCTGGAGATCTACTCAGTCTCAGGACTCTTAGAGAAAGCGCTGT 606  
QY 241 GCAGCGGCTCTCTGGGGATCTGAGCGCTCAGGAGTATGATCAGTACTCTGGTCTTGG 300  
|||||

Db 607 GCANGCGGCTTCTGGGGATCTGAGCGCTCANGACGTGATNAGTACTCTGGTCTGG 666  
RESULT 8  
HS1100H13/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
COMMENT  
On Dec 5, 2000 this sequence version replaced gi:10198625.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/projects/Celegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 20, constructed by the Sanger Centre Chromosome 20  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr20  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. RP5-1100H13 is from  
the library RPCI-5 constructed by the group of Pieter de Jong. For  
further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2  
This sequence is the entire insert of clone RP5-1100H13 The true  
left end of clone RP11-12201 is at 47827 in this sequence. The true  
right end of clone RP5-927M24 is at 16404 in this sequence.  
FEATURES  
Location/Qualifiers  
1..116792  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="R2PD:RPCIP704H131100"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="q11.2"  
/clone="RP5-1100H13"  
/clone\_11b="RPCI-5"  
1331..1465  
/note="MER58C repeat: matches 68..89 of consensus"  
1587..13349  
/gene="dJ1100H13.1"  
Join(<1587..1723,4379..4487,5360..5338,8641..8789,  
9265..13349)  
/gene="dJ1100H13.1"  
/product="dJ1100H13.1.1 (KIAA1219 (similar to Drosophila  
GH9358 and C. elegans D2085.5) isoform 1)"  
/note="match: cDNAs: Em:AB033045 Em:AL132998"



Db 49509 TGAGAGTAAACCCAGCGGACTTCTCACAAGCGCTCAACACAGCTACAAACT-C 49451  
QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 240  
Db 49450 GCACAAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 49391  
QY 241 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGACTCAGGCTGTCTGG 300  
Db 49390 GCAGGCGGCTTGTGGGGATGTGA-CGCTCAGGAGCTGATGAGGACTCAGGCTGTCTGG 49332

RESULT 9  
AC009068/c  
LOCUS AC009068 Homo sapiens chromosome 16 clone RP11-314K3, linear PRI 27-APR-2001  
DEFINITION Homo sapiens chromosome 16 clone RP11-314K3, complete sequence.  
AC009068  
VERSION AC009068.10 GI:13811892  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 160356)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 160356)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 160356)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (27-APR-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Apr 27, 2001 this sequence version replaced gi:12000278.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence  
Estimated Total Number of Errors is 0.3.

FEATURES  
source  
1. .160356  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-314K3"  
BASE COUNT 37469 a 43134 c 43308 g 36445 t  
ORIGIN

Query Match 67.0%; Score 201; DB 9; Length 160356.  
Best Local Similarity 100.0%; Pred. No. 3.9e-43;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AGCTGGATACCTTACCAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 159  
Db 50029 AGCTGGATACCTTACCAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 49970  
QY 160 TCAACACATGTACAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 219  
Db 49969 TCAACACATGTACAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 49910  
QY 220 ACTTCTAGAGAAAGGCGCTGTGTCAGGCGGCTTCTGCGGGATGTGAGCGCTCAGGAGCTG 279  
Db 49909 ACTTCTAGAGAAAGGCGCTGTGTCAGGCGGCTTCTGCGGGATGTGAGCGCTCAGGAGCTG 49850

Chromosome 16

QY 280 ATGAGGTACTCGTGGTCTGG 300  
Db 49849 ATGAGGTACTCGTGGTCTGG 49829

RESULT 10  
AC123908/c  
LOCUS AC123908 Homo sapiens chromosome 16 clone CTD-2542L18, linear PRI 28-MAR-2003  
DEFINITION Homo sapiens chromosome 16 clone CTD-2542L18, complete sequence.  
AC123908  
VERSION AC123908.3 GI:29336199  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 185321)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 185321)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (05-JUN-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 185321)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (06-SEP-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 185321)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Mar 28, 2003 this sequence version replaced gi:22748376.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence  
Estimated Total Number of Errors is 0.1.

FEATURES  
source  
1. .185321  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="CTD-2542L18"  
BASE COUNT 40674 a 50904 c 51869 g 41874 t  
ORIGIN

Query Match 67.0%; Score 201; DB 9; Length 185321;  
Best Local Similarity 100.0%; Pred. No. 3.8e-43;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 AGCTGGATACCTTACCAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 159  
Db 95218 AGCTGGATACCTTACCAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 95159  
QY 160 TCAACACATGTACAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 219  
Db 95158 TCAACACATGTACAACTCCGACCACTCCAGCTCTGGAGAGTACTCAGTCTCAGG 95099  
QY 220 ACTTCTAGAGAAAGGCGCTGTGTCAGGCGGCTTCTGCGGGATGTGAGCGCTCAGGAGCTG 279  
Db 95098 ACTTCTAGAGAAAGGCGCTGTGTCAGGCGGCTTCTGCGGGATGTGAGCGCTCAGGAGCTG 95039  
QY 280 ATGAGGTACTCGTGGTCTGG 300

Db 95038 ATGAGTACTCGGTCTGG 95018

RESULT 11

AC137771

LOCUS

DEFINITION Homo sapiens chromosome 16 clone CTD-2139A24, WORKING DRAFT

AC137771

SEQUENCE, 45 unordered pieces.

AC137771

AC137771.1 GI:26006513

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 245210)

TITLE DOE Joint Genome Institute.

JOURNAL Sequencing of Human Chromosome 16

REFERENCE 2 (bases 1 to 245210)

TITLE Direct Submission

JOURNAL DOE Joint Genome Institute.

REFERENCE Submitted (03-DEC-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

AUTHORS Center: Joint Genome Institute

TITLE Center Code: JGI

JOURNAL Web site: <http://www.jgi.doe.gov>

COMMENT

Project Information

Center Project Name: 680949

Center clone name: CITB-HL\_2139A24

-----

Summary Statistics

Consensus quality: 221831 bases at least Q40

Consensus quality: 226530 bases at least Q30

Consensus quality: 229658 bases at least Q20

Estimated insert size: 170000; agarose-fp estimation

Estimated insert size: 240810; sum-of-contigs estimation

Quality coverage: 17.31 in Q20 bases; agarose-fp estimation

Quality coverage: 12.22 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 45 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1084: contig of 1084 bp in length

\* 1085 1184: gap of unknown length

\* 1185 2248: contig of 1064 bp in length

\* 2249 2348: gap of unknown length

\* 2349 3435: contig of 1087 bp in length

\* 3436 3535: gap of unknown length

\* 3536 4637: contig of 1102 bp in length

\* 4638 4737: gap of unknown length

\* 4738 5816: contig of 1079 bp in length

\* 5817 5916: gap of unknown length

\* 5917 6951: contig of 1035 bp in length

\* 6952 7051: gap of unknown length

\* 7052 8599: contig of 1548 bp in length

\* 8600 8699: gap of unknown length

\* 8700 10508: contig of 1809 bp in length

\* 10509 10609: gap of unknown length

\* 10609 11838: contig of 1130 bp in length

\* 11739 11839: gap of unknown length

\* 13754 13754: contig of 1916 bp in length

\* 13755 13855: gap of unknown length

\* 13855 14996: contig of 1142 bp in length

\* 14997 15096: gap of unknown length

\* 15097 16852: contig of 1756 bp in length

\* 16853 16952: gap of unknown length

18502: contig of 1550 bp in length

18503 18602: gap of unknown length

18603 20736: contig of 2134 bp in length

20737 20836: gap of unknown length

20837 2243: contig of 1407 bp in length

2244 22343: gap of unknown length

22344 24310: contig of 1967 bp in length

24311 24410: gap of unknown length

24411 26675: contig of 2265 bp in length

26676 26775: gap of unknown length

26776 28257: contig of 1482 bp in length

28258 28357: gap of unknown length

28358 29668: contig of 1311 bp in length

29669 29768: gap of unknown length

29769 31503: contig of 1735 bp in length

31504 31603: gap of unknown length

31604 33962: contig of 2359 bp in length

33963 34062: gap of unknown length

34063 36099: contig of 2037 bp in length

36100 38832: contig of 2633 bp in length

38833 38932: gap of unknown length

38933 41464: contig of 2532 bp in length

41465 41564: gap of unknown length

41565 44085: contig of 2521 bp in length

44086 44185: gap of unknown length

44186 47028: contig of 2843 bp in length

47029 47128: gap of unknown length

47129 49395: contig of 2267 bp in length

49396 49495: gap of unknown length

49496 51435: contig of 1940 bp in length

51436 51535: gap of unknown length

51536 55706: contig of 4171 bp in length

55707 55806: gap of unknown length

55807 60092: contig of 4286 bp in length

60093 60192: gap of unknown length

60193 63448: contig of 3156 bp in length

63449 66966: contig of 3518 bp in length

66967 71746: gap of unknown length

67067 71746: contig of 4680 bp in length

71747 71846: gap of unknown length

71847 74430: contig of 2584 bp in length

74431 74530: gap of unknown length

74531 79515: contig of 4985 bp in length

79516 79616: gap of unknown length

79616 83377: contig of 3762 bp in length

83378 83477: gap of unknown length

83478 88208: contig of 4731 bp in length

88209 88308: gap of unknown length

88309 93320: contig of 5012 bp in length

93321 93420: gap of unknown length

93421 97345: contig of 3925 bp in length

97346 97445: gap of unknown length

97446 100893: contig of 3448 bp in length

100894 100993: gap of unknown length

100994 107097: contig of 6104 bp in length

107098 107197: gap of unknown length

107198 111963: contig of 4766 bp in length

111964 112063: gap of unknown length

112064 116267: contig of 4204 bp in length

116268 116367: gap of unknown length

116368 123219: contig of 6852 bp in length

123220 123319: gap of unknown length

123320 245210: contig of 121891 bp in length.

Location/Qualifiers

1..245210

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="16"

/clone="CTD-2139A24"

/clone\_lib="Caltech human BAC library D"

FEATURES

SOURCE



BASE COUNT 62851 a 55678 c 56421 g 65631 t 4629 others  
ORIGIN

Query Match 67.0%; Score 201; DB 2; Length 245210;

Best Local Similarity 100.0%; Pred. No. 3.6e-43;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 ACCTGGATTAACCTGGAGATCAACACAGCGGGACTTCCCTCACACAGCGC 159  
|||||

Db 115926 ACCTGGATTAACCTGGAGATCAACACAGCGGGACTTCCCTCACACAGCGC 115985  
|||||

QY 160 TCAACACATGTACAACTCCGACGACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGG 219  
|||||

Db 115986 TCAACACATGTACAACTCCGACGACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGG 116045  
|||||

QY 220 ACTTCTAGAGAAAGCGCTGGTGCAGCGGCTTGGTGGGGGATGTGAGCGCTCAGGACGTG 279  
|||||

Db 116046 ACTTCTAGAGAAAGCGCTGGTGCAGCGGCTTGGTGGGGGATGTGAGCGCTCAGGACGTG 116105  
|||||

QY 280 ATGAGGTACTCGTGGTCTGG 300  
|||||

Db 116106 ATGAGGTACTCGTGGTCTGG 116126

# RESULT 12

AK091519

LOCUS

DEFINITION Homo sapiens cDNA FLJ34200 fis, clone FCBF3019663.

ACCESSION AK091519

VERSION AK091519.1 GI:21749909

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,

Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,

Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,

Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,

Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,

Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K.,

Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M.,

Takahashi-Fuji,I., A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,

Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 3097)

Isogai,T. and Yamamoto,J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan

(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

Location/Qualifiers

1. .3097

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="FCBBF3019663"

/tissue\_type="brain"

/clone\_lib="FCBBF3"

/dev\_stage="fetal"

/note="cloning vector: pME18SFL3"

726 a 703 c 786 g 882 t

BASE COUNT

ORIGIN

# Query Match

Best Local Similarity 99.3%; Pred. No. 1.8e-26;

Matches 151; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 60  
|||||

Db 303 ACGAATCCGGACCTGGTCAAGGATATGTGGGACACTCGTATAGCCAACTCCGAGTGT 362  
|||||

QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGACATGCCAAGCTGGGATAACTTGCACCTTGA 120  
|||||

Db 363 CTGCTGACAGCTTTGTGAGACAGCAGGACATGCCAAGCTGGGATAACTTGCACCTTGA 421  
|||||

QY 121 TGGAGATCAACACAGCGGGACTTTCTCTACA 152  
|||||

Db 422 TGGAGATCAACACAGCGGGACTTTCTCTACA 453  
|||||

# RESULT 14

AB097169

LOCUS

DEFINITION Xenopus laevis Psf2 gene, complete cds.

ACCESSION AB097169

VERSION AB097169.1 GI:29365480

KEYWORDS Xenopus laevis (African clawed frog)

ORGANISM	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Xenopus; Xenopodinae; Xenopus.
REFERENCE 1	
AUTHORS	Kubota,Y., Takase,Y., Komori,Y., Hashimoto,Y., Arata,T., Kamimura,Y., Araki,H. and Takisawa,H.
TITLE	A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication
JOURNAL	Genes Dev. 17 (9), 1141-1152 (2003)
MEDLINE	22615695
PUBMED	12730133
REFERENCE 2	(bases 1 to 933)
AUTHORS	Kubota,Y. and Takisawa,H.
TITLE	Direct Submission
JOURNAL	Submitted (28-NOV-2002) Yumiko Kubota, Osaka University, Department of Biology, Graduate School of Science, Machikaneyamacho 1-1, Toyonaka, Osaka 560-0043, Japan
	(E-mail:ykubota@bio.sci.osaka-u.ac.jp, Tel:81-6-6850-5554, Fax:81-6-6850-5554)
FEATURES	Location/Qualifiers
source	1..933
	/organism="Xenopus laevis"
	/mol_type="genomic DNA"
gene	/db_xref="taxon:8355"
	21..578
CDS	/gene="Psf2"
	21..578
	/gene="Psf2"
	/function="essential for DNA replication"
	/codon_start=1
	/product="Psf2"
	/protein_id="BAC66459.1"
	/db_xref="GI:29365481"
	/translation="MDASEVEFLAEKEQTVIPNFSLDKVLIGDLPFNPSLPVEV PLWALNQRKRCRLVPPEDMDVEKLEAIRDQREETFTPMSPYYMELTKLLNH AADINPKADEIRTLKVDITLWDTRIAKLRLSADSFVKQGEAHKLDNLTLMEINTIGTFF TESLNHYKLRITSLONPPEGQSDY"
BASE COUNT	282 a 200 c 208 g 243 t
ORIGIN	
Query Match	41.9%; Score 125.8; DB 5; Length 933;
Best Local Similarity	68.1%; Pred. No. 9.2e-23;
Matches 175; Conservative	0; Mismatches 82; Indels 0; Gaps 0;
QY	1 ACAGAAATCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTCT 60
Db	352 ATGAGATTGGTACACTGTAAAGACACCTGGGATCAAGAATAGCAAACTCGGCTCT 411
QY	61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCATGCCAAGCTGGATAACTTGACCTTGA 120
Db	412 CTGCTCACAGCTTTGTGAGGGGCAGGAAGCTCATGCCAAGCTGGATAACCTGACGCTAA 471
QY	121 TGGAGATCAACACCGCGGACTTTCCTCACAGAAGCGCTCAACCATGTACAACTCC 180
Db	472 TGGAAATTAAACCATTCGGAACATTTTTTACTGAGTCTTTAAACCATGTACAAGCTAC 531
QY	181 GCACGAACCTCAGCCTCTGGAGAGTACTCACTCTCAGGACTTCAGAGAAAGCGCTGT 240
Db	532 GTACCGACCTGCGAAGACCCAGAGGAGGACGATCAGAGTACTTAAGCAATTTGTCGTGA 591
QY	241 GCAGCGCGCTTGCTGGG 257
Db	592 ATTATCTGGCTGCTGAG 608
RESULT 15	
AC117085/c	
LOCUS	AC117085 207307 bp DNA linear HTG 13-MAY-2003
DEFINITION	Rattus norvegicus clone CH230-247G13, WORKING DRAFT SEQUENCE.
ACCESSION	AC117085
VERSION	AC117085.8 GI:30579062
KEYWORDS	HTG; HTGS-PHASE2; HTGS-DRAFT; HTGS FULLTOP.

SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
REFERENCE	1 (bases 1 to 207307)
AUTHORS	Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S,, Adams,C,, Alder,J., Allen,C, Allen,H,, Albrooks,S,, Amin,A,, Anguiano,D., Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H., Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F., Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Claveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., DeSouza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dlinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,I., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Fregegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunarathne,P., Haaland,W., Hamil,C., Hamilton,C., Hamillon,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,D., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louissegh,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naif,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelmech,O., Okwuonu,G., Olarunpungoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puzao,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Tjebbes,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,Y., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 207307)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 207307)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 13, 2003 this sequence version replaced gi:24942311. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas ( <a href="http://www.hgsc.bmc.tmc.edu/projects/rat/">http://www.hgsc.bmc.tmc.edu/projects/rat/</a> ). Each contig described

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information

```
Center project name: GMGE
Center clone name: CH230-247G13
----- Summary Statistics
```

Assembly program: Atlas 3.0;  
Consensus quality: 198554 bases at least Q40  
Consensus quality: 200661 bases at least Q30  
Consensus quality: 201900 bases at least Q20  
Estimated insert size: 210626; sum-of-contrigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contrigs estimation

```

* NOTE: Estimated insert size may differ from sequence length .
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 207307: contig of 207307 bp in length.
*

```

## FEATURES

```

1. 207307
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/Clone="CH230-247G13"
87039..177061
/note="clone_boundary
clone_end:Sp6
site:MboI
end_sequence:B2169551"
150202..150357
/note="clone_boundary
clone_end:T7
site:MboI

```

## misc\_feature

```
87039. .177061
/note="clone_boundary
clone_end:Sp6
site:MboI
```

## misc\_feature

```

/note="clone_boundary
clone_end:T7
site:Mbol

```

```

BASE COUNT 43941 a 54080 c 55681 g 48886 t 4719 others
ORIGIN
end_sequence: 64706350

```

	Query Match	37.3%	Score 111.8	DB 2	Length 207307
	Best Local Similarity	78.4%	Pred. No. 1.9e-19		
	Matches 134	Conservative 0	Mismatches 37	Indels 0	Gaps 0
QY	100	AGCTGGATACCTTGACCTTGGATGAGATCAACACAGCGGGACATTTCTCTACACAAGCGC	159		
DB	196735	AGCTGGACAACCTTGACCTTGGATGAGATCAACACAGCTGGGGCCCTTCTCTACCCCAAGCTC	196676		
QY	160	TCAACACATGTTACAAACTCGGCACGAACTCCAGCCCTCTGGAGAGTACTCAGTCTTCAGG	219		
DB	196675	TCAACACATGTTACAAACTCGGCACGAACTTCAGCCCTCTGAGAGCTCAGAGTCTCAGG	196616		
QY	220	ACTTCTAGAGAAAGCCCTGGTGTGAGCGGGCTTGCTGGGGGATGTGAGCGCT	270		
DB	196615	ACTTCTAGCCAGAGGCGCTCTGGTGTGCAAAACCCCTTGTGGGTTAGAGTGAT	196565		



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 11:32:16 ; Search time 1272.67 Seconds  
(without alignments)  
5729.195 Million cell updates/sec

Title: US-09-854-124-5

Perfect score: 300

Sequence: 1 acgaataccgaccctgttc.....tgaggtactcgtggtctgg 300

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estinu:\*
  - 4: em\_estnu:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_htc:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_hic:\*
  - 12: gb\_est3:\*
  - 13: gb\_est4:\*
  - 14: gb\_est5:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_nam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rod:\*
  - 26: em\_gss\_phg:\*
  - 27: em\_gss\_vrl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	300	100.0	533	14	CB112523 K-EST0154
C 2	300	100.0	559	12	BM126492 lf06h06.y
C 3	300	100.0	659	13	BU729618 01-E-COI-
C 4	300	100.0	716	10	BE382866 601297755

5	300	100.0	801	12	BI196248
6	300	100.0	876	10	BG181130
7	300	100.0	896	10	BE562088
8	300	100.0	914	10	BE547290
9	300	100.0	943	10	BF683514
10	300	100.0	970	9	AL560669
11	300	100.0	1007	12	BM449472
12	300	100.0	1064	13	BQ277667
13	300	100.0	1071	13	BU184963
14	300	100.0	1201	9	AL560880
15	300	100.0	1201	9	AL582250
16	299.6	99.9	1201	9	AL582217
17	299	99.7	810	12	BQ009838
18	298.8	99.6	1193	9	AL529785
19	298.4	99.5	488	14	CB160336
20	298.4	99.5	696	9	AW249012
21	298.4	99.5	712	13	BU625683
22	298.4	99.5	714	10	BG104289
23	298.4	99.5	716	12	BG825252
24	298.4	99.5	730	10	BE296429
25	298.4	99.5	790	10	BE795306
26	298.4	99.5	866	10	BF795157
27	298.4	99.5	867	13	BQ233393
28	298.4	99.5	891	11	BC022839
29	298.4	99.5	921	10	BE799212
30	298.4	99.5	932	10	BE796384
31	298.4	99.5	1028	10	BE561044
32	298.4	99.5	1030	10	BF311745
33	298.4	99.5	1182	11	AF125098
34	298	99.3	833	13	BU597296
35	297.4	99.1	494	14	H64860
36	296.8	98.9	895	13	BU539659
37	293.2	97.7	1056	13	BA456910
38	292.4	97.5	667	9	AU126087
39	292	97.3	966	9	AL582077
40	291.4	97.1	1201	9	AL562756
41	290	96.7	849	10	BE514071
42	289.6	96.5	1007	10	BE795838
43	288.6	96.2	1201	9	AL526847
44	288	96.0	761	12	BI257993
45	287.4	95.8	966	10	BE791539

ALIGNMENTS

RESULT 1  
CB112523/c  
LOCUS K-EST0154352 L6ChoCK0 Homo sapiens cDNA clone L6ChoCK0-4-D04 5',  
DEFINITION mRNA sequence.  
ACCESSION CB112523  
VERSION CB112523.1 GI:27938330  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 533)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 4 row: D column: 04  
High quality sequence stop: 533.

FEATURES  
source

Location/Qualifiers  
1. .533  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L6choCK0-4-D04"  
/sex="M"  
/cell\_line="Cho-CK"  
/lab\_host="Top10F"  
/clone\_lib="L6choCK0"

Note: Organ: Liver; Vector: pCNS-D2; Site1: EcoRI;  
Site2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library.

BASE COUNT 109 a 143 c 135 g 146 t  
ORIGIN  
Query Match 100.0%; Score 300; DB 14; Length 533;  
Best Local Similarity 100.0%; Pred. No. 1.8e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCGAGTGT 60  
DB 309 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCGAGTGT 250  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 120  
DB 249 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 190  
QY 121 TGGAGATCAACACACAGCGGGACTTTCCTCACACAAGCGCTCAACACATGTCACAACTCC 180  
DB 189 TGGAGATCAACACACAGCGGGACTTTCCTCACACAAGCGCTCAACACATGTCACAACTCC 130  
QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
DB 129 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 70  
QY 241 GCAGCGGCTGCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 300  
DB 69 GCAGCGGCTGCTGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 10

## RESULT 2

LOCUS BM126492 559 bp mRNA linear EST 12-MAR-2002  
DEFINITION if06h06.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens  
cDNA clone IMAGE:5675770 5' similar to TR:Q9Y248 Q9Y248 HSPC037  
PROTEIN. i, mRNA sequence.

ACCESSION BM126492  
VERSION BM126492.1 GI:17121044

## KEYWORDS

SOURCE EST.

## ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 559)  
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hallier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas  
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.

TITLE  
JOURNAL  
COMMENT

Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished  
Other ESTs: if06h06.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brownjefas.harvard.edu) This sequence now available from the IMAGE  
consortium, for clone orders contact: info@image.llnl.gov  
High quality sequence stop: 437.

FEATURES  
source

Location/Qualifiers  
1. .559  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5675770"  
/sex="Both"  
/tissue\_type="Islets of Langerhans"  
/dev\_stage="Adult"  
/lab\_host="DH10B"

/clone\_lib="Melton Normalized Human Islet 4 N4-HIS 1"  
/note="Organ: Pancreas; Vector: pSPORL1; Site1: Not 1;  
Site2: Sal 1; Starting library constructed using  
SuperScript Plasmid Library kit (Life Technologies). cDNA  
made by oligo-dT priming. Size-selected by column  
fractionation: average insert size 1.08 kb. Library was  
amplified once on solid support and plasmid DNA from  
library was prepared. The library DNA was normalized by  
method #4 from Bonaldo, Lennon, and Soares 1996 Genome  
Research 6:791-806; 0.5 microgram single-stranded library  
plasmid DNA was mixed with 5 micrograms PCR product  
representing library inserts and hybridized to an EcoT of  
20. Single-stranded (unhybridized) plasmids were isolated  
by hydroxyapatite chromatography and used to make this  
library."

BASE COUNT 157 a 131 c 142 g 129 t

## ORIGIN

Query Match 100.0%; Score 300; DB 12; Length 559;  
Best Local Similarity 100.0%; Pred. No. 1.8e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCGAGTGT 60  
DB 118 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCGAGTGT 177

QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 120  
DB 178 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTGACCTTGA 237  
QY 121 TGGAGATCAACACACAGCGGGACTTTCCTCACACAAGCGCTCAACACATGTCACAACTCC 180  
DB 238 TGGAGATCAACACACAGCGGGACTTTCCTCACACAAGCGCTCAACACATGTCACAACTCC 297  
QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
DB 298 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 357  
QY 241 GCAGCGGCTGCTGGGGATGTGAGCGCTCAGACGTGATGAGGTACTCGTGGTTCTGG 300  
DB 358 GCAGCGGCTGCTGGGGATGTGAGCGCTCAGACGTGATGAGGTACTCGTGGTTCTGG 417

## RESULT 3

BU729618/c

LOCUS BU729618 659 bp mRNA linear EST 09-OCT-2002  
 DEFINITION UI-E-CQ1-ay-b-18-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone  
 ACCESSION BU729618  
 VERSION BU729618  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 659)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 FORWARD  
 POLYA-Yes.

## FEATURES

Location/Qualifiers  
 1..659  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-CQ1-ay-b-18-0-UI"  
 /tissue\_type="optic nerve"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-CQ1"  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-E-CQ1 is a normalized cDNA library containing the  
 following tissue(s): optic nerve. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (AT)18 tail. The sequence tag for this library is  
 CCATTAAGTG. This library was created for the program, Gene  
 Discovery in the Visual System, supported by National Eye  
 Institute (NEI).  
 TAG\_LIB=UI-E-CQ1  
 TAG\_TISSUE=human optic nerve  
 TAG\_SEQ=CCATTAAGTG"

BASE COUNT 149 a 159 c 151 g 200 t  
 ORIGIN

Query Match 100.0%; Score 300; DB 13; Length 659;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db 509 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 450  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 120

Db 449 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 390  
 QY 121 TGGAGATCAACACACAGCGGACTTCTCCTCACACAAGCGCTCAACACACATGTACAAACTCC 180  
 Db 389 TGGAGATCAACACACAGCGGACTTCTCCTCACACAAGCGCTCAACACACATGTACAAACTCC 330  
 QY 181 GCACGAACTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 Db 329 GCACGAACTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 270  
 QY 241 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300  
 Db 269 GCAGGCGGCTTGTGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 210

## RESULT 4

BE382866 716 bp mRNA linear EST 21-JUL-2000  
 LOCUS 601297755F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3627852 5',  
 mRNA sequence.  
 BE382866  
 ACCESSION BE382866.1 GI:9328231  
 VERSION  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 716)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLCM311 row: n column: 13  
 High quality sequence stop: 651.

## FEATURES

Location/Qualifiers  
 1..716  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3627852"  
 /tissue\_type="neuroblastoma"  
 /clone\_lib="NIH\_MGC\_19"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

BASE COUNT 185 a 179 c 195 g 157 t  
 ORIGIN

Query Match 100.0%; Score 300; DB 10; Length 716;  
 Best Local Similarity 100.0%; Pred. No. 2e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 Db 337 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 396  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 120  
 Db 397 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 456

QY 121 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 180  
 |||||  
 Db 457 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 516  
 |||||  
 QY 181 GCAGCAACCTCCAGCCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGCCGTGGT 240  
 |||||  
 Db 517 GCAGCAACCTCCAGCCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGCCGTGGT 576  
 |||||  
 QY 241 GCAGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTGTGTTCTGG 300  
 |||||  
 Db 577 GCAGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTGTGTTCTGG 636  
 |||||

RESULT 5  
 B1196248  
 LOCUS 801 bp mRNA linear EST 10-JUL-2001  
 DEFINITION 602754709F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4890147 5',  
 mRNA sequence.  
 ACCESSION B1196248  
 VERSION B1196248.1 GI:14651268  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE NIH-MGC http://mgc.nci.nih.gov/.  
 AUTHORS 1 (bases 1 to 801)  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LICM1767 row: b column: 04  
 High quality sequence stop: 781.  
 FEATURES  
 Location/Qualifiers  
 1..801  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4890147"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; CDNA made by oligo-dt priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 208 a 196 c 216 g 181 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 12; Length 801;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 |||||  
 Db 356 ACAGAAATCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 415  
 |||||  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGCCTTGA 120  
 |||||  
 Db 416 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGCCTTGA 475  
 |||||  
 QY 121 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 180

Db 476 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 535  
 |||||  
 QY 181 GCAGCAACCTCCAGCCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGCCGTGGT 240  
 |||||  
 Db 536 GCAGCAACCTCCAGCCCTCTGGAGAGTACTAGTCTCAGGACTTCTAGAGAAAGCCGTGGT 595  
 |||||  
 QY 241 GCAGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTGTGTTCTGG 300  
 |||||  
 Db 596 GCAGCGGCTTGCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCTGTGTTCTGG 655  
 |||||

RESULT 6  
 B181130  
 LOCUS 876 bp mRNA linear EST 06-FEB-2001  
 DEFINITION 602329264F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4430823 5',  
 mRNA sequence.  
 ACCESSION B181130  
 VERSION B181130.1 GI:12687833  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE NIH-MGC http://mgc.nci.nih.gov/.  
 AUTHORS 1 (bases 1 to 876)  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10184 Row: O Column: 16  
 High quality sequence stop: 674.  
 FEATURES  
 Location/Qualifiers  
 1..876  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4430823"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-dt primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 246 a 217 c 214 g 199 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 10; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAAATCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 |||||  
 Db 115 ACAGAAATCCGACCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 174  
 |||||  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGCCTTGA 120  
 |||||  
 Db 175 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAAGCTGGATAACTTGCCTTGA 234  
 |||||  
 QY 121 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 180  
 |||||  
 Db 235 TGGAGATCAACACAGCGGGGACTTTCCTCACACAAGCGCTCAACACATGTACAAACTCC 294  
 |||||  
 QY 181 GCAGCAACCTCCAGCCCTGGTCAAGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 240



```

|||||
Db      295 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGGT 354
QY      241 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 300
Db      355 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 414

RESULT 7
LOCUS      BE562088
DEFINITION 601345038F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678080 5',
mRNA sequence.
ACCESSION BE562088
VERSION   BE562088.1 GI:9805808
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 896)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCN354 row: k column: 09
High quality sequence start: 24
High quality sequence stop: 840.
FEATURES             Location/Qualifiers
     source
     1..896
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3678080"
     /tissue_type="Burkitt lymphoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_8"
     /note="Organ: lymph; Vector: pOTB7; Site:1: XhoI; Site:2:
     EcoRI; cDNA made by oligo-dT priming. Directionally
     cloned into EcoRI/XhoI sites using the following 5'
     adaptor: GGCACGAG(G). Size-selected >500bp for average
     insert size 1.8kb. Library constructed by Ling Hong in
     the laboratory of Gerald M. Rubin (University of
     California, Berkeley) using ZAP-cDNA synthesis kit
     (Stratagene) and Superscript II RT (Life Technologies)."
```

*Lymph*

```

Db      618 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGGT 677
QY      241 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 300
Db      678 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 737

RESULT 8
LOCUS      BE547290
DEFINITION 601073715F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3460010 5',
mRNA sequence.
ACCESSION BE547290
VERSION   BE547290.1 GI:9775935
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 914)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8453 row: m column: 03
High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source
     1..914
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3460010"
     /tissue_type="cervical carcinoma cell line"
     /lab_host="DH10B"
     /clone_lib="NIH_MGC_12"
     /note="Organ: cervix; Vector: pCMV-SPORT6; Site:1: NotI;
     Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 1.4 kb. Library prepared by Life
     Technologies."
```

*Cervical*  
*CA cell line*

```

BASE COUNT      254 a 227 c 238 g 195 t
ORIGIN

Query Match      100.0%; Score 300; DB 10; Length 914;
Best Local Similarity 100.0%; Pred. No. 2.3e-72;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60
Db      104 ACAGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 163
QY      61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 120
Db      164 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGATACCTTGACCTTGA 223
QY      121 TGGAGATCAACACCGAGCGGACTTTCTCCACACAGCGCTCAACACACATGTACAAACTCC 180
Db      224 TGGAGATCAACACCGAGCGGACTTTCTCCACACAGCGCTCAACACACATGTACAAACTCC 283
QY      181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGGT 240
Db      284 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGGT 343
QY      241 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 300
Db      344 GCAGGCGGCTTCTGGGGGATGTGAGCGCTCAGGACGCTGATGAGGTACTCGTGGTCTCGG 403
```

RESULT 9  
 LOCUS BF683514  
 DEFINITION 602139714Fl NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4300947 5',  
 mRNA sequence.  
 ACCESSION BF683514  
 VERSION BF683514.1 GI:11968922  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 943)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHCMI160 row: 1 column: 04  
 High quality sequence stop: 709.

FEATURES  
 source  
 1. 943  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4300947"  
 /tissue\_type="leiomyosarcoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH-MGC\_46"  
 /note="Organ: uterus; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACACGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 214 a 253 c 272 g 204 t

Query Match. 100.0%; Score 300; DB 10; Length 943;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 DB 369 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 428  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCGCAAGCTGGATTAACCTTGACCTTGA 120  
 DB 429 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCGCAAGCTGGATTAACCTTGACCTTGA 488  
 QY 121 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACATGTACAAATCC 180  
 DB 489 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACATGTACAAATCC 548  
 QY 181 GCAGAACCTCCAGCCCTCGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 549 GCAGAACCTCCAGCCCTCGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 608  
 QY 241 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 300  
 DB 609 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 668

RESULT 10  
 LOCUS AL560669  
 DEFINITION AL560669 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 Homo sapiens cDNA clone CS0DL003YI21 5-PRIME, mRNA sequence.  
 ACCESSION AL560669  
 VERSION AL560669.2 GI:31284799  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 970)  
 Li.W.B., Gruber.C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 On Feb 15, 2001 this sequence version replaced gi:12907354.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 7987.r For  
 more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DL003AE11Q1&cluster=7987.r>. Contact :  
 Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DL003AE11Q1.

FEATURES  
 source  
 1. 970  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DL003YI21"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 257 a 243 c 251 g 219 t

Query Match. 100.0%; Score 300; DB 9; Length 970;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-72;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 DB 419 ACGAAATCCGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 478  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCGCAAGCTGGATTAACCTTGACCTTGA 120  
 DB 479 CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCGCAAGCTGGATTAACCTTGACCTTGA 538  
 QY 121 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACATGTACAAATCC 180  
 DB 539 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACATGTACAAATCC 598  
 QY 181 GCAGAACCTCCAGCCCTCGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 599 GCAGAACCTCCAGCCCTCGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 658  
 QY 241 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 300  
 DB 659 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTCTGG 718

```
RESULT 11
BM449472      1007 bp      mRNA      linear      EST 05-FEB-2002
LOCUS         5', mRNA sequence.
DEFINITION    AGENCOURT_6400913 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493679
ACCESSION     BM449472
VERSION       BM449472
KEYWORDS      EST.
SOURCE        BM449472.1 GI:18498512
ORGANISM      Homo sapiens (human)
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL       1 (bases 1 to 1007)
COMMENT       NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM12117 row: m column: 08
              High quality sequence stop: 711.
              Location/Qualifiers
                1..1007
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5493679"
                  /tissue_type="retinoblastoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_67"
                  /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                  Average insert size 1.75 kb. Library constructed by Life
                  Technologies."
              BASE COUNT      259 a      254 c      262 g      231 t      1 others
              ORIGIN
                Query Match      100.0%; Score 300; DB 12; Length 1007;
                Best Local Similarity 100.0%; Pred. No. 2.4e-72;
                Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

                QY      1  ACGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60
                Db      1  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      61  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTTGACCTTGA 120
                Db      61  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      417  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTTGACCTTGA 476
                Db      417  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      121  TGGAGATCAACACACAGCGGGACTTTCTCACACAAGCGCTCAACACATGTACAACTCC 180
                Db      121  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      477  TGGAGATCAACACACAGCGGGACTTTCTCACACAAGCGCTCAACACATGTACAACTCC 536
                Db      477  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      181  GCAGCAACCTCCAGCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240
                Db      181  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      537  GCAGCAACCTCCAGCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 596
                Db      537  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      241  GCAGCGCGCTTGCCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 300
                Db      241  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      597  GCAGCGCGCTTGCCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 656
                Db      597  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 12
BQ277667
LOCUS         BQ277667
DEFINITION    AGENCOURT_7048107 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5804514
              5', mRNA sequence.
              BQ277667      1064 bp      mRNA      linear      EST 07-MAY-2002
              AGENCOURT_7048107 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5804514
              5', mRNA sequence.
```

```
ACCESSION     BQ277667
VERSION       BQ277667.1 GI:20487875
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 1064)
              NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM2042 row: d column: 19
              High quality sequence stop: 675.
              Location/Qualifiers
                1..1064
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5804514"
                  /tissue_type="teratocarcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_109"
                  /note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
                  XhoI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGCACGAG(G). Library constructed by Ling Hong in the
                  laboratory of Gerald M. Rubin (University of California,
                  Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                  Superscript II RT (Life Technologies). Note: this is a
                  NIH_MGC Library."
              BASE COUNT      275 a      268 c      273 g      245 t      3 others
              ORIGIN
                Query Match      100.0%; Score 300; DB 13; Length 1064;
                Best Local Similarity 100.0%; Pred. No. 2.5e-72;
                Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

                QY      1  ACGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60
                Db      1  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      364  ACGAAATCCGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 423
                Db      364  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      61  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTTGACCTTGA 120
                Db      61  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      424  CTGCTGACAGCTTTGTGAGACAGCAGGAGGCACATGCCAAGCTGGATAACTTTGACCTTGA 483
                Db      424  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      121  TGGAGATCAACACACAGCGGGACTTTCTCACACAAGCGCTCAACACATGTACAACTCC 180
                Db      121  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      484  TGGAGATCAACACACAGCGGGACTTTCTCACACAAGCGCTCAACACATGTACAACTCC 543
                Db      484  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      181  GCAGCAACCTCCAGCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240
                Db      181  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      544  GCAGCAACCTCCAGCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 603
                Db      544  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      241  GCAGCGCGCTTGCCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 300
                Db      241  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                QY      604  GCAGCGCGCTTGCCTGGGGGATGTGAGCGCTCAGGACGTGATGAGGTACTCGTGGTTCTGG 663
                Db      604  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 13
BUI84963
LOCUS         BUI84963
DEFINITION    AGENCOURT_7186943 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6002100
              5', mRNA sequence.
              BUI84963      1071 bp      mRNA      linear      EST 04-SEP-2002
              AGENCOURT_7186943 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6002100
              5', mRNA sequence.
```

```

VERSION      BU184963.1  GI:22698947
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens

REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: CGAP (Stanford)
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.lnl.gov
              Plate: LNCM2269 row: m column: 13
              High quality sequence stop: 718.
              Location/Qualifiers
                1..1071
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6002100"
                  /tissue_type="hepatocellular carcinoma, cell line"
                  /lab_host="PH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_100"
                  /note="Organ: liver; Vector: pOPB7; Site:1; XhoI; Site:2;
                  EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGCACGAG(G). Size-selected >500bp for average insert size
                  1.8kb. Library constructed by Ling Hong in the laboratory
                  of Gerald M. Rubin (University of California, Berkeley)
                  using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                  II RT (Life Technologies). Note: this is a NIH_MGC
                  Library."
              BASE COUNT      283 a 262 c 275 g 246 t 5 others
              ORIGIN

              Query Match      100.0%; Score 300; DB 13; Length 1071;
              Best Local Similarity 100.0%; Pred. No. 2.5e-72;
              Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

              QY 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60
                  |||||||
              Db 359 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 418
                  |||||||

              QY 61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGGATACTTGACCTTGA 120
                  |||||||
              Db 419 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGGATACTTGACCTTGA 478
                  |||||||

              QY 121 TGGAGATCAACACACAGCGGACTTTCCCTCACAAAGCGCTCAACACATGTACAAACTCC 180
                  |||||||
              Db 479 TGGAGATCAACACACAGCGGACTTTCCCTCACAAAGCGCTCAACACATGTACAAACTCC 538
                  |||||||

              QY 181 GCACGAACTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCGTGT 240
                  |||||||
              Db 539 GCACGAACTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCGTGT 598
                  |||||||

              QY 241 GCAGGCGGCTTCTGCTGGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300
                  |||||||
              Db 599 GCAGGCGGCTTCTGCTGGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 658
                  |||||||

              RESULT 14
              LOCUS      AL560880
              DEFINITION  AL560880 Homo sapiens B CELLS (RAMOS CELL LINE) EST 31-MAY-2003
              ACCESSION  AL560880 Homo sapiens cDNA clone CSODL005YK19 5-PRIME, mRNA sequence.
              VERSION    AL560880
              KEYWORDS   EST.

              AL560880.2  GI:31285009
              EST.
              SOURCE       Homo sapiens (human)
              ORGANISM     Homo sapiens

              REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              TITLE        Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
              JOURNAL      Full-length cDNA libraries and normalization
              COMMENT      Unpublished
              On Feb 15, 2001 this sequence version replaced gi:12907768.
              Contact: Genoscope
              Genoscope - Centre National de Sequencage
              Bp 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 7987.r For
              more information about this cluster, see
              http://www.genoscope.cns.fr/
              cgi-bin/cluster.cgi?seq=CSODL005AF100P1&cluster=7987.r. Contact :
              Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CSODL005AF100P1.
              Location/Qualifiers
                1..1201
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="CSODL005YK19"
                  /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                  /cell_line="RAMOS CELL LINE"
                  /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                  25-NORMALIZED"
                  /note="1st strand cDNA was primed with a NotI-oligo(dT)
                  primer. Five prime end enriched, double-strand cDNA was
                  digested with Not I and cloned into the Not I and EcoR V
                  sites of the pCMVSPORT 6 vector. Library was normalized."
              BASE COUNT      322 a 267 c 292 g 283 t 37 others
              ORIGIN

              Query Match      100.0%; Score 300; DB 9; Length 1201;
              Best Local Similarity 100.0%; Pred. No. 2.6e-72;
              Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

              QY 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 60
                  |||||||
              Db 426 ACGAATCCGGACCCCTGGTCAAGGATATGTGGACACTCGTATAGCCAAACTCCGAGTGT 485
                  |||||||

              QY 61 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGGATACTTGACCTTGA 120
                  |||||||
              Db 486 CTGCTGACAGCTTTGTGACACAGCAGGAGGCACATGCCAAGCTGGGATACTTGACCTTGA 545
                  |||||||

              QY 121 TGGAGATCAACACAGCGGAGCTTTCCCTCACAAAGCGCTCAACACATGTACAAACTCC 180
                  |||||||
              Db 546 TGGAGATCAACACAGCGGAGCTTTCCCTCACAAAGCGCTCAACACATGTACAAACTCC 605
                  |||||||

              QY 181 GCACGAACTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCGTGT 240
                  |||||||
              Db 606 GCACGAACTCCAGCTCTGGAGAGTACTCAGTCTCAGGACTCTAGAGAAAGGCGTGT 665
                  |||||||

              QY 241 GCAGGCGGCTTCTGCTGGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 300
                  |||||||
              Db 666 GCAGGCGGCTTCTGCTGGGGGATGTGAGCGCTCAGGAGCTGATGAGGTACTCGTGGTCTGG 725
                  |||||||

              RESULT 15
              LOCUS      AL582250/c
              DEFINITION  AL582250 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
              ACCESSION  AL582250 Homo sapiens cDNA clone CSODL005YF11 3-PRIME, mRNA sequence.
              VERSION    AL582250
              KEYWORDS   EST.
              AL582250.2  GI:31320467
              EST.

```

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
AUTHORS Li, W. B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT On Feb 16, 2001 this sequence version replaced gi:12950048.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 7987.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DL005CC06NP1&cluster=7987.r. Contact :  
Feng Liang Email: fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DL005CC06NP1.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DL005YF11"  
/cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
/cell\_line="RAMOS CELL LINE"  
/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 291 a 238 c 281 g 293 t 98 others  
ORIGIN  
Query Match 100.0%; Score 300; DB 9; Length 1201;  
Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 60  
Db |||||  
QY 746 ACGAATCCGGACCCCTGGTCAAGGATATGTGGGACACTCGTATAGCCAAACTCCGAGTGT 687  
Db |||||  
QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGACATGCCAAGCTGGGATAACTTGACCTTGA 120  
Db |||||  
QY 686 CTGCTGACAGCTTTGTGAGACAGCAGGACATGCCAAGCTGGGATAACTTGACCTTGA 627  
Db |||||  
QY 121 TGGAGATCAACACACCGGGGACTTTCTCTACACAAGCGCTCAACCACTGTACAAACTCC 180  
Db |||||  
QY 626 TGGAGATCAACACACCGGGGACTTTCTCTACACAAGCGCTCAACCACTGTACAAACTCC 567  
Db |||||  
QY 181 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 240  
Db |||||  
QY 566 GCACGAACCTCCAGCCTCTGGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGGCGCTGT 507  
Db |||||  
QY 241 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGAGTGTGAGTACTCTGTTCTG 300  
Db |||||  
QY 506 GCAGGCGGCTTGTCTGGGGGATGTGAGCGCTCAGGAGTGTGAGTACTCTGTTCTG 447  
Search completed: October 14, 2003, 14:09:03  
Job time : 1279.67 secs

**THIS PAGE BLANK (USPTO)**

JOURNAL Patent: JP 2002191363-A 10946 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002191363-A/10946  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
 PI SAITO,  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 PI KEIICHI NAGAI, TETSUJI OTSUKI  
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC  
 10,  
 PC C12P21/02, C12P1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH key  
 Location/Qualifiers  
 FT CDS (37)..  
 TURES Location/Qualifiers  
 source 1..858  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 227 a 214 c 222 g 195 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 6; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-69;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACAGAAATCCGACCCCTGTCAGGACATATGCGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 DB 368 ACAGAAATCCGACCCCTGTCAGGACATATGCGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCGGACACTCGGACCTTGAACCTTGA 120  
 DB 428 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCGGACACTCGGACCTTGAACCTTGA 120  
 QY 121 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACCATGTACAACTCC 180  
 DB 488 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACCATGTACAACTCC 180  
 QY 181 GCAGCAACCTCCAGCCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 548 GCAGCAACCTCCAGCCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 QY 241 GCAGCGGGCTTGTGGGGATGTGAGCGCTCAGGACCTGATGAGTACTCGTGGTCTGG 300  
 DB 608 GCAGCGGGCTTGTGGGGATGTGAGCGCTCAGGACCTGATGAGTACTCGTGGTCTGG 300  
 RESULT 2  
 AK001275 858 bp mRNA linear PRI 01-AUG-2002  
 LOCUS Homo sapiens cDNA FLJ10413 fls, clone NT2RP1000063.  
 DEFINITION AK001275  
 ACCESSION AK001275.1 GI:7022428  
 VERSION oligo capping; fls (full insert sequence).  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
 Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
 Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
 Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
 Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 REFERENCE 2 (bases 1 to 858)  
 AUTHORS Isogai, T. and Otsuki, T.

Direct Submission  
 Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction, 5'- & 3'-end one pass sequencing and clone selection;  
 Helix Research Institute (supported by Japan Key Technology Center  
 etc.) and Department of Virology, Institute of Medical Science,  
 University of Tokyo.  
 Location/Qualifiers  
 source 1..858  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP1000063"  
 /cell\_line="NT2"  
 /cell\_type="teratocarcinoma"  
 /clone\_lib="NT2RP1"  
 /note="cloning vector: pUC19FL3-mRNA from NT2 neuronal  
 precursor cells after 48-hours retinoic acid (RA)  
 induction."  
 CDS  
 37..594  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAA91595.1"  
 /db\_xref="GI:7022429"  
 /translation="MDAAVEFLAEKELVLIIPNFSLDKLYLGGDLGPNPGLPVEV  
 PLWAIINLKKORCKRLPPPEMDVDEKLEKWRDHERKEETTFPMPSPYIMETKLLNH  
 ASDNIPKADEIRTLVDMNDRTRAKLRVSDSFVRQOEHAHAKLDNLMLBINTSGTFL  
 TQALNHMYKLRTHLPUESTQSQDF"  
 BASE COUNT 227 a 214 c 222 g 195 t  
 ORIGIN  
 Query Match 100.0%; Score 300; DB 9; Length 858;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-69;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ACAGAAATCCGACCCCTGTCAGGACATATGCGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 DB 368 ACAGAAATCCGACCCCTGTCAGGACATATGCGGACACTCGTATAGCCAAACTCCGAGTGT 60  
 QY 61 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCGGACACTCGGACCTTGAACCTTGA 120  
 DB 428 CTGCTGACAGCTTTGTGAGACAGCAGGAGGACATGCGGACACTCGGACCTTGAACCTTGA 120  
 QY 121 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACCATGTACAACTCC 180  
 DB 488 TGGAGATCAACACAGCGGGACTTTCCTCACAAAGCGCTCAACACCATGTACAACTCC 180  
 QY 181 GCAGCAACCTCCAGCCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 DB 548 GCAGCAACCTCCAGCCCTCTGAGAGTACTCAGTCTCAGGACTTCTAGAGAAAGCCCTGGT 240  
 QY 241 GCAGCGGGCTTGTGGGGATGTGAGCGCTCAGGACCTGATGAGTACTCGTGGTCTGG 300  
 DB 608 GCAGCGGGCTTGTGGGGATGTGAGCGCTCAGGACCTGATGAGTACTCGTGGTCTGG 300  
 RESULT 3  
 AF151880 1174 bp mRNA linear PRI 18-MAY-2000  
 LOCUS Homo sapiens CGI-122 protein mRNA, complete cds.  
 DEFINITION AF151880  
 ACCESSION AF151880  
 VERSION AF151880.1 GI:4929712  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1174)





**TITLE** Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics

**JOURNAL** Genome Res. 10 (5), 703-713 (2000)

**MEDLINE** 20272150

**PUBMED** 10810093

**REFERENCE** 2 (bases 1 to 1174)

**AUTHORS** Lin, W.-C.

**TITLE** Direct Submission

**JOURNAL** Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan

**FEATURES** Location/Qualifiers

source 1..1174

1..1174 /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="16"

77..634

/codon\_start=1

/product="CGI-122 protein"

/protein\_id="AA034117.1"

/db\_xref="GI:4929713"

/translation="MDAAEVEFLAEKELVTIIPNFSLDKIYLGDDLPNPGLPVEV PLWAIILKQKCRLLPPEWMDVEKLEMRDHERKEETFTPMSPYMETLKLLNH ASDNIPKADEIRTLVKDMDTRIAKLRVSADSFVRQEAHAKLDNLTLMEINTSGTFL TOALNHYKLRNLQPLESTQSDF"

BASE COUNT 300 a 299 c 307 g 268 t

ORIGIN

Query Match 100.0%; Score 300; DB 9; Length 1174;

Best Local Similarity 100.0%; Pred. No. 4.5e-69;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAAATCCGACCCCTGCTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 60

DB 408 ACGAAATCCGACCCCTGCTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 467

QY 61 CTGCTGACAGCTTTGTGAGACAGGAGGACATGCCAAGCTGGATTAACCTTGACTTGA 120

DB 468 CTGCTGACAGCTTTGTGAGACAGGAGGACATGCCAAGCTGGATTAACCTTGACTTGA 527

QY 121 TGGAGATCAACACACCGGGGACTTTCTTCCACACAAAGCGCTCAACACATGTACAACTCC 180

DB 528 TGGAGATCAACACACCGGGGACTTTCTTCCACACAAAGCGCTCAACACATGTACAACTCC 587

QY 181 GCAGCAACCTCCAGCCTCTGGAGACTCTCAGTCTCAGGACTCTAGACAAAGGCTGTGT 240

DB 588 GCAGCAACCTCCAGCCTCTGGAGACTCTCAGTCTCAGGACTCTAGACAAAGGCTGTGT 547

QY 241 GCAGCGGCTTGTCTGGGGATGTAGCGCTCAGGAGTGTATGAGTACTCGTGTCTGTG 300

DB 648 GCAGCGGCTTGTCTGGGGATGTAGCGCTCAGGAGTGTATGAGTACTCGTGTCTGTG 707

**RESULT 4**

LOCUS BC010164 1180 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, HSPC037 protein, clone MGC:19836 IMAGE:4098007, mRNA, complete cds.

ACCESSION BC010164

VERSION BC010164.1 GI:14603431

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1180)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**AUTHORS** Strausberg, R.

**TITLE** Direct Submission

**JOURNAL** Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
nisc\_mgc@hgr.nih.gov  
Contact: Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjong, S.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 27 Row: 1 Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction, similarity but not identity to protein.

**FEATURES** Location/Qualifiers

source 1..1180

1..1180 /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:19836 IMAGE:4098007"

/tissue\_type="Brain, neuroblastoma"

/clone\_lib="NIH\_MGC\_19"

/lab\_host="DH10B-R"

/note="Vector: pONB7"

68..625

/codon\_start=1

/product="HSPC037 protein"

/protein\_id="AAH010164.1"

/db\_xref="GI:14603432"

/translation="MDAAEVEFLAEKELVTIIPNFSLDKIYLGDDLPNPGLPVEV PLWAIILKQKCRLLPPEWMDVEKLEMRDHERKEETFTPMSPYMETLKLLNH ASDNIPKADEIRTLVKDMDTRIAKLRVSADSFVRQEAHAKLDNLTLMEINTSGTFL TOALNHYKLRNLQPLESTQSDF"

BASE COUNT 316 a 292 c 307 g 265 t

ORIGIN

Query Match 100.0%; Score 300; DB 9; Length 1180;

Best Local Similarity 100.0%; Pred. No. 4.5e-69;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGAAATCCGACCCCTGCTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 60

DB 399 ACGAAATCCGACCCCTGCTCAAGGATATGTGGACACTCGTATAGCCAACTCCGAGTGT 458

QY 61 CTGCTGACAGCTTTGTGAGACAGGAGGACATGCCAAGCTGGATTAACCTTGACTTGA 120

DB 459 CTGCTGACAGCTTTGTGAGACAGGAGGACATGCCAAGCTGGATTAACCTTGACTTGA 518

QY 121 TGGAGATCAACACACCGGGGACTTTCTTCCACACAAAGCGCTCAACACATGTACAACTCC 180

DB 519 TGGAGATCAACACACCGGGGACTTTCTTCCACACAAAGCGCTCAACACATGTACAACTCC 578

QY 181 GCAGCAACCTCCAGCCTCTGGAGACTCTCAGTCTCAGGACTCTAGACAAAGGCTGTGT 240

DB 579 GCAGCAACCTCCAGCCTCTGGAGACTCTCAGTCTCAGGACTCTAGACAAAGGCTGTGT 638

QY 241 GCAGCGGCTTGTCTGGGGATGTAGCGCTCAGGAGTGTATGAGTACTCGTGTCTGTG 300

DB 639 GCAGCGGCTTGTCTGGGGATGTAGCGCTCAGGAGTGTATGAGTACTCGTGTCTGTG 698

**RESULT 5**

**THIS PAGE BLANK (USPTO)**

8-17-2